ADVANCES IN PHYSICS

A QUARTERLY SUPPLEMENT of the PHILOSOPHICAL MAGAZINE

VOLUME 3 JULY 1954 NUMBER 11

PRICE per part 15*s.* 0*d.*
PRICE per annum £2 15*s.* 0*d.* post free

PRINTED AND PUBLISHED BY TAYLOR & FRANCIS LTD.
RED LION COURT, FLEET ST., LONDON E.C.4

Early Scientific Publications

DIARY OF ROBERT HOOKE, M.A., M.D., F.R.S.
1672-1680

Edited by **H. W. ROBINSON** and **W. ADAMS**
Recommended for publication by the Royal Society, London

25/-
net

"This vivid record of the scientific, artistic and social activities of a remarkable man during remarkable years has too long remained in obscurity."—Extract from foreword by Sir Frederick Gowland Hopkins, O.M., President of the Royal Society.

MATHEMATICAL WORK OF JOHN WALLIS, D.D., F.R.S.

By **J. F. SCOTT, Ph.D., B.A.**

12/6
net

"His work will be indispensable to those interested in the early history of The Royal Society. I commend to all students of the Seventeenth Century, whether scientific or humane, this learned and lucid book."—Extract from foreword by Prof. E. N. da C. Andrade, D.Sc., Ph.D., F.R.S.
Recommended for publication by University of London

CORRESPONDENCE AND PAPERS OF EDMOND HALLEY

Arranged and Edited by **EUGENE FAIRFIELD MACPIKE**

21/-
net

First published on behalf of The History of Science Society by Oxford University Press. Now re-issued by Taylor & Francis, Ltd.

MEMOIRS OF SIR ISAAC NEWTON'S LIFE

5/-
net

By **WILLIAM STUKELEY, M.D., F.R.S., 175**
From an Original Manuscript
Now in the possession of the Royal Society, London

HEVELIUS, FLAMSTEED AND HALLEY

12/6
net

Three Contemporary Astronomers and their Mutual Relations
By **EUGENE FAIRFIELD MACPIKE**
Published by arrangement with The History of Science Society

Established over 150 years

TAYLOR & FRANCIS, LTD.
RED LION COURT, FLEET ST., LONDON E.C.
PRINTERS & PUBLISHERS OF SCIENTIFIC BOOKS

CONTENTS

Annealing of Cold Worked Metals By Paul A. Beck, University of Illinois, Urbana, U.S.A 245

Electrons in Lattice Fields. By H. Fröhlich, Department of Theoretical Physics, University of Liverpool . . . 325

ADVANCES IN PHYSICS

A QUARTERLY SUPPLEMENT

of the

PHILOSOPHICAL MAGAZINE

VOLUME 3 JULY 1954 NUMBER 11

Annealing of Cold Worked Metals

By PAUL A. BECK
University of Illinois, Urbana, U.S.A.

§ 1. INTRODUCTION

THE investigations of Taylor and Quinney[1], Suzuki[2], Clarebrough, Hargreaves, Michell and West[3], Bever and Ticknor[4], Gordon[5] and others clearly established that of the order of 1% to 10% of the energy expended in plastic deformation remains stored in cold deformed metals. For instance in heavily cold worked copper the stored energy of cold work may amount to approximately 30 calories per mole, or of the order of 1% of the heat of fusion. Even though the stored energy of cold work is relatively small, it is an indication of the instability of the deformed state. As a result of this instability, the deformed metal has a tendency to return to the low energy soft-annealed condition, when the temperature is high enough for sufficient atomic mobility. The present paper deals with the various processes whereby the well-annealed condition may be approached, namely, recovery, subgrain growth* and recrystallization. Earlier work in the whole area has been reviewed extensively by Burgers[6] and Barrett[7]; other reviews cover more specifically the field of subgrains[8] or of recrystallization and grain growth[9, 10, 11, 12, 13]. The present paper aims primarily at the presentation of developments resulting from the manifold experimental research activities of the last few years.

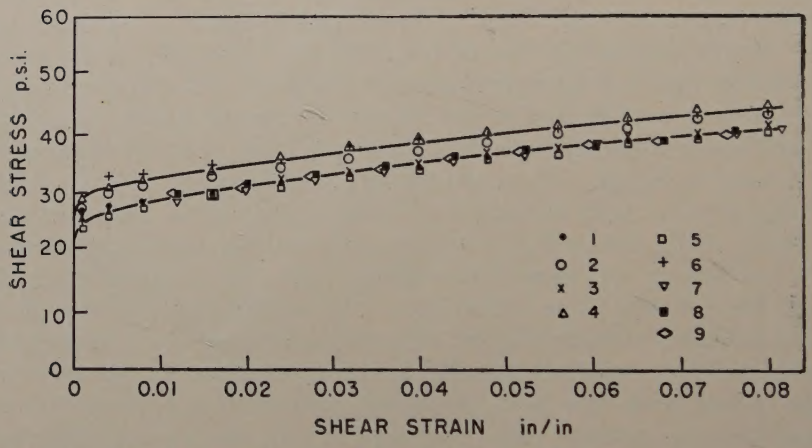
* Some investigators prefer to include the phenomenon of subgrain growth (or ' polygonization ') under the heading of recovery. However, true recovery phenomena, characterized by their kinetics (figs. 2, 6, 7, and 9), do occur without any interface migration on a scale detectable by present methods. Although the kinetics of subgrain growth is not well known as yet, there is some indication that it may be different from that of true recovery. Also, subgrain growth, as detected by present microscopic or x-ray diffraction methods, apparently requires distinctly higher annealing temperatures (or longer times) than recovery. In view of the facts mentioned, the separation of these phenomena into two distinct classes appears to be well justified. Unfortunately, the usage of nomenclature in this field is rather confused ; some authors use the term ' recovery ' in such a broad sense that even recrystallization phenomena are included.

§ 2. Recovery

2.1. *Recovery of Work Hardening*

In single crystals of Zn, when deformed in such a manner that only one glide system operates (' easy glide '), and no bending is involved, the work hardening, as measured by the increase of the resolved shear stress on the operating slip system, quickly decreases when the deformed crystal is stored at room temperature. This phenomenon of recovery (' Kristallerholung ') has been first described by Polanyi and Schmid[14] and later studied in more detail by Haase and Schmid[15], who showed that in extended Zn single crystals it gradually proceeded to completion, so that the soft-annealed condition was in effect reached within a period of approximately one day at room temperature. Even after repeated cycles of stressing and recovery, provided that the deformation was glide on a

Fig. 1



Nine successive shear stress–strain curves from the same pure zinc single crystal, showing complete recovery by intermediate annealing treatments of 1 hour at 260°c. [Ref. (16).]

single slip system without bending, recovery went to completion at room temperature in about one day, so that the stress–strain curve upon every new load application was identical with the original one. Similar results with Zn single crystals were recently obtained by Li, Washburn and Parker[16], who repeated the cycle of shear deformation and complete recovery in one hour at 260°c as many as nine times. It is characteristic of these experiments that after recovery anneal the well annealed single crystal condition was again attained, as shown for instance by the return to the original flow stress values (fig. 1). No interfaces of any kind were introduced into the single crystal either during the deformation process or the subsequent annealing. As pointed out particularly by Masing[17], the process of recovery is homogeneous on a microscopic scale, and it does

not involve any interface migration. Although one may not as yet be certain as to the details of the exact nature of the lattice imperfections that are responsible for work hardening in pure shear on a single slip system, it is clear that these imperfections must be of an easily recoverable type. It is significant that under such conditions of deformation Honeycombe[18] found no x-ray asterism in single crystals of Cd, even after 100% extension, or more. If the deformation does involve bending and x-ray asterism is observed, the resulting work hardening is greater, but only a fraction of this work hardening is recoverable[18]. This is particularly true for polycrystalline metals, where the constraints of neighbour grains force each crystal to deform in a complex manner. The point is well illustrated by the results of Tietz, Anderson and Dorn[25] with polycrystalline high purity aluminium specimens deformed in tension at three different temperatures to the same flow stress of 12 300 p.s.i., as measured at 78°K. The recovery of the flow stress after various periods of annealing at 305°K was measured at 78°K. The extent of recovery at 305°K was approximately 5·7%, 10% and 14% of the total work hardening for the specimens deformed at 273°, 194° and 78°K, respectively. These results show that the easily recoverable component represents a variable fraction of the constant total work hardening, and that this fraction decreases with increasing temperature of deformation. By extrapolation, one may expect that plastic deformation of polycrystalline high purity aluminium at higher temperatures, such as 60°C or 100°C, may result in work hardening with no recoverable component at all.

Cherian, Pietrokowsky and Dorn[31] found that polycrystalline tension specimens of commercial purity Al after small tension strains at room temperature show very marked recovery of the yield stress on annealing at temperatures up to 100°C, without any lowering of the flow stress at higher strains (' metarecovery '). They reported that annealing in the temperature range of about 150°–200°C results, not only in a lowering of the yield stress, but also in a lowering of the flow stress at high strains (' orthorecovery '). The two phenomena appear to have different kinetics. It is to be noted, however, that in their study of the recovery of zinc single crystals deformed in pure shear, Parker and his associates did not observe metarecovery. Also, there appears to be no indication of the occurrence of metarecovery in the study by Dorn and his associates of recovery in pure Al[25]. It is very likely that the occurrence of ' metarecovery ', observed with Al of commercial purity, is not a general phenomenon, but that it is associated with the impurities in the particular metal tested by Cherian, Pietrokowsky and Dorn.

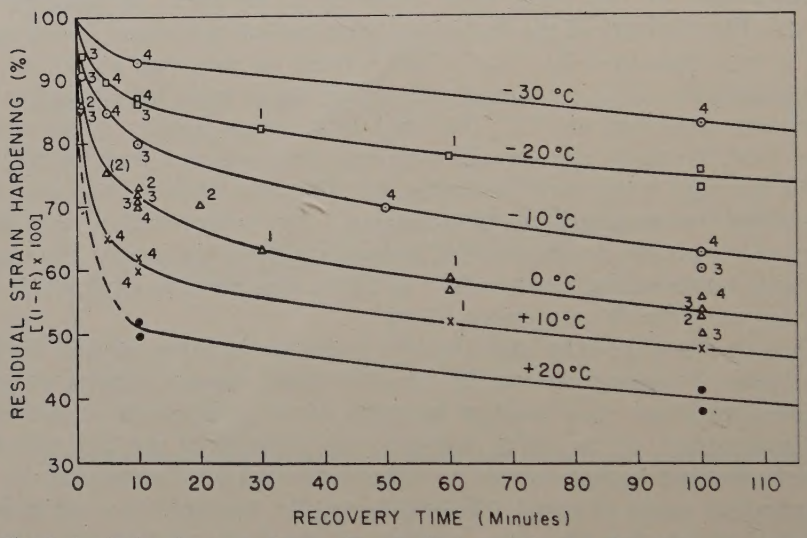
2.2. *Kinetics of the Recovery of Work Hardening*

Drouard, Washburn and Parker[19] measured isothermal recovery at a series of different temperatures in zinc single crystals deformed at 78°K in pure shear along a single glide system, without bending. The amount of recovery was measured by the decrease of the flow stress in percents of the

difference in flow stress between the deformed and the fully annealed condition. Isothermal curves for various temperatures are shown in fig. 2. The data clearly show the type of kinetics typical of recovery: at all temperatures the rate of the process is highest initially, and then it decreases with continued recovery. It is to be noted that the kinetics of recovery is significantly different from that of a process involving nucleation and interface migration[20], where the rate is at first small, then it increases to a maximum and decreases again (see fig. 42). This difference further supports the conclusion in the previous paragraph, that recovery does not involve the migration of interfaces.

The isothermal kinetics of recovery is shown more fully in fig. 3, where the recovery, expressed in percents of the total recoverable work hardening,

Fig. 2

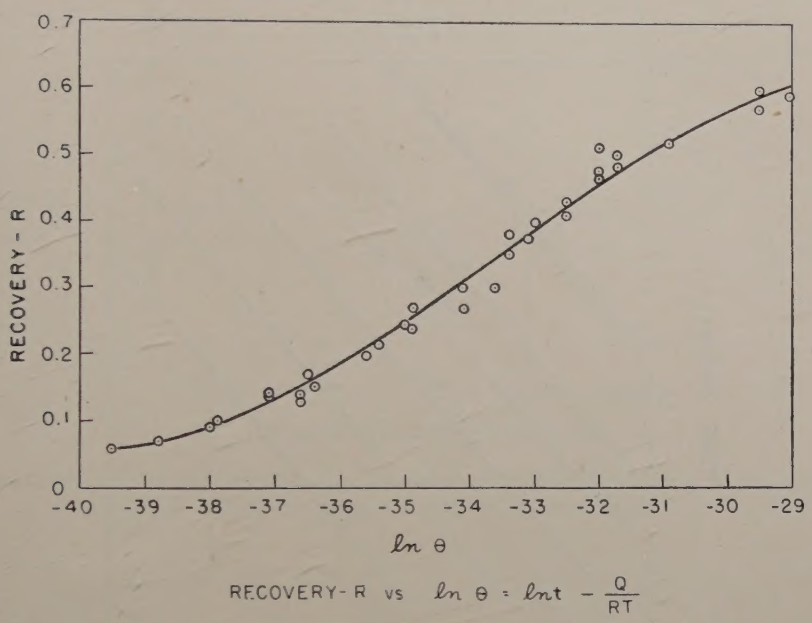


Residual strain hardening of pure Zn single crystals as a function of annealing time at various temperatures, after 8% shear strain at −50°C. [Ref. (19).]

is plotted against the logarithm of the temperature-adjusted recovery time, combining all of Drouard, Washburn and Parker's data for various temperatures by means of a single activation energy. The central part of the curve has an inflection point and it may be considered as a straight line over a range corresponding to a time ratio of about 250. At short annealing times the experimental recovery values are larger and at long annealing times they are smaller than those extrapolated from the straight line relationship. Very similar kinetics were found previously by Masing and Raffelsieper[21] for pure Al single crystals deformed in tension at room temperature. They considered the logarithmic recovery kinetics (corresponding to the straight line central portion in fig. 3) as a confirmation of the theory of recovery by Kuhlmann, Masing and Raffelsieper[22]. As shown by Kuhlmann[23], the assumption that the activation energy for

recovery decreases linearly with progressive recovery leads directly to a linear relationship between recovery and the logarithm of the annealing time. The deviations from this relationship at very short and very long annealing times are also accounted for by the theory of Kuhlmann, Masing and Raffelsieper.

Fig. 3



Fraction of strain hardening recovered (R) as a function of the logarithm ot the temperature-adjusted annealing time, from the data in fig. 2. [Ref. (19).]

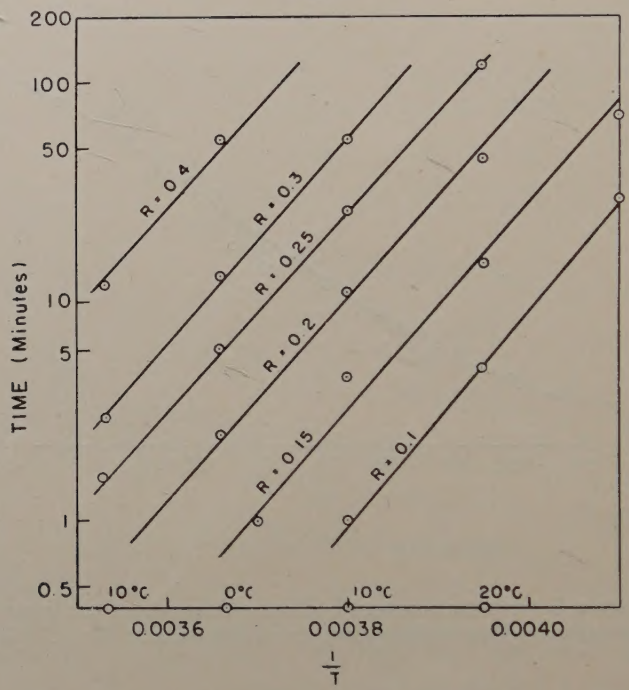
2.3. *Activation Energy of the Recovery of Work Hardening*

According to the results of Drouard, Washburn and Parker[19], the temperature dependence of the rate of recovery of Zn single crystals sheared at 78°K and annealed in the temperature range of −20°C to +10°C can be described satisfactorily by means of a single activation energy of 20 kilocalories/mole (fig. 4). This value is in good agreement with that for the self-diffusion in Zn crystals along the hexagonal axis[24], suggesting that the elementary process in recovery may be similar to that in diffusion. However, the independence of the activation energy from the extent of recovery in these experiments appears to contradict the theory of Kuhlmann, Masing and Raffelsieper. The experimental data of Raffelsieper with Al, which indicated a linear decrease of the activation energy with progressive recovery (fig. 5), were obtained under less favourable conditions than those of Drouard, Washburn, and Parker, since in the extended Al single crystals a substantial portion of the work hardening was not recoverable.

2.4. *Recovery of the Stored Energy of Cold Work*

It is significant that the kinetics of the low temperature (60° and 100°C) isothermal release of the stored energy of cold work from cold rolled polycrystalline Cu and Al, according to Borelius, Berglund and Sjoberg[26] also obeys a logarithmic law, since the rate of evolution of heat varies inversely with the annealing time (see fig. 6 and eqn. (1)).

Fig. 4

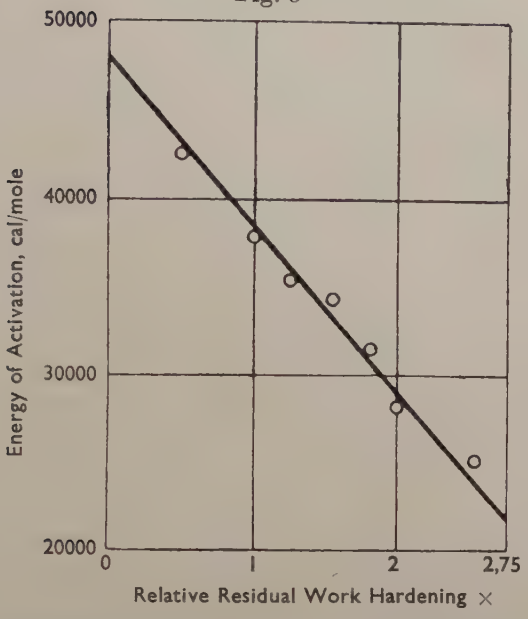


Temperature dependence of the annealing time required for various fractions of recovery in sheared zinc single crystals (data from fig. 2). The slope of the lines corresponds to a constant activation energy of 20 kcal per mole. [Ref. (19).]

$$P = \frac{RT}{b} \cdot \frac{1}{t} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

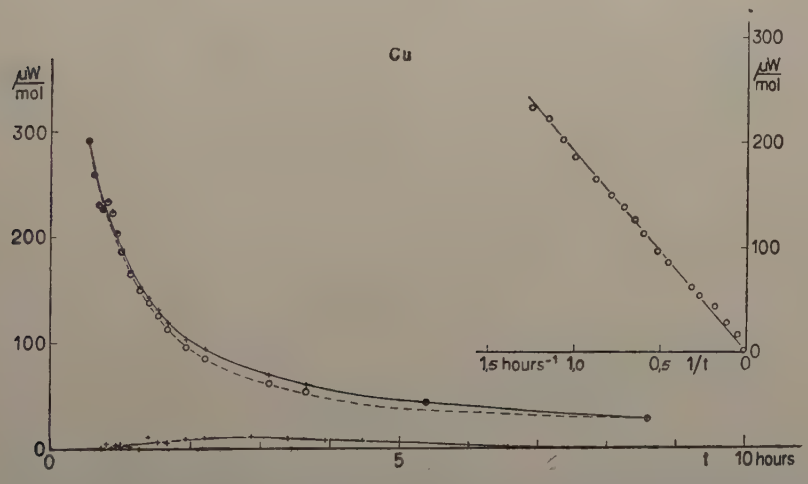
where P is the rate of energy release at the time t. This result establishes a close analogy between the energy release and the recovery of flow stress, as discussed above, and suggests that the reaction associated with the energy release studied by Borelius and his associates is also recovery, not involving the migration of interfaces—a very probable inference, considering the low temperatures and the short annealing periods involved. It would be interesting to have flow stress and calorimetric recovery data for identical specimens, in order to ascertain whether or not the time constants for the two recovery curves are the same. Further questions that arise in this connection concern the effect of the temperature of

Fig. 5



Decrease of activation energy for recovery of strain hardening in Al single crystals with progressive recovery. [Ref. (22).]

Fig. 6



Rate of release of stored energy of cold work at 100°c from Cu rolled to double length, as a function of annealing time, and in inset reciprocal annealing time. [Ref. (26).]

deformation on the release of the stored energy of cold work. In analogy with the effect of temperature on the recovery of yield stress, one may expect, for instance, that with decreasing deformation temperature the rate of energy release and also the total energy released during recovery increase. Also, it is quite possible that plastic deformation at somewhat higher temperatures (60°–100°C for Al) may not lead to the storage of any recoverable energy, so that the energy stored under such conditions could be released only in recrystallization.

2.5. *Recovery of X-ray Line Broadening*

The kinetics of the isothermal sharpening of x-ray diffraction lines broadened by cold working has been studied by Van Arkel and Burgers[27] in the case of tungsten. They found that in cold drawn thoriated tungsten wires considerable line sharpening takes place prior to recrystallization, and that the kinetics of this process is qualitatively similar to that illustrated above for recovery in the case of flow stress and energy release. It has been held by some investigators that the recovery of line broadening prior to recrystallization occurs only in metals of high melting point, such as Mo and W. In other metals, for instance in cold worked Cu, Ni, alpha brass, and steel, Wilson and Thomassen[28] found that line sharpening on annealing occurs simultaneously with softening. Averbach[29] found that in brass filings line sharpening was a result of recrystallization, rather than of a recovery process prior to and distinct from recrystallization. However, a recent study of isothermal line sharpening in cold rolled pure Al single crystals by Lutts and Beck[3] gave definite separation between line sharpening and the major portion of softening (fig. 24); it also indicated clearly a typical recovery-type kinetics for line sharpening. The kinetics itself is, of course, not sufficient to prove that the observed line sharpening in effect takes place by a recovery process in the strict sense, without sub-boundary migration.

Undoubtedly, 'recovery' of line broadening prior to recrystallization may occur in low-melting metals, as well as in those of high melting point. Whether or not it may be actually observed under certain conditions apparently depends mainly on the relative rates of recovery and of recrystallization. In the particular experiments with Al, referred to above, recrystallization, which presumably depends on the availability of 'nuclei' in suitable orientations, was delayed by using single crystals of the (110) [112] orientation and by careful techniques, eliminating contamination of the resulting single orientation rolling texture by crystal fragments of other orientations. Recovery, which does not involve nucleation or interface migration, is less likely to be affected by such procedures. The relative rates of recovery and of recrystallization depend also on the presence of finely dispersed particles of a second phase. This may be seen for instance in the results of Van Arkel and Burgers[27], who found much more distinct separation of the two phenomena in thoriated tungsten than in pure tungsten. The finely dispersed thoria particles slow down the migration of grain boundaries necessary for

recrystallization, while their effect on the rate of recovery is apparently smaller, if any. In view of the results with brass mentioned above, solid solution of Zn in Cu apparently does not tend to delay recrystallization more than recovery.

Just as the recoverable component of work hardening, the line broadening in the cold worked state also increases with decreasing deformation temperature, according to the investigations of Wilson and Thomassen[31] and of Paterson and Orowan[32]. It is a reasonable assumption that all of the line broadening is, at least potentially, recoverable (may be annealed out by a recovery mechanism, provided that recrystallization does not set in prematurely). The results of Paterson and Orowan indicate that the variation of the amount of line broadening with the deformation temperature cannot be explained on the basis of increasing recovery at higher deformation temperatures, but that it must be considered as a true effect of the temperature of deformation. It is therefore an interesting question whether or not the much larger line broadening observed in W, as compared with Al, may be accounted for by the higher rate of recovery at the deformation temperature in the lower melting metal, as suggested by Williamson and Hall[33].

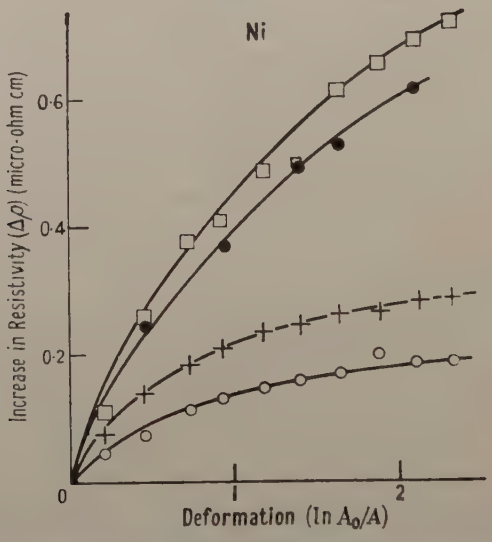
2.6. *Recovery of Electrical Resistivity*

In view of the recent publication of an extensive review by Broom[78], the present remarks will be restricted to a brief summary of newer results. Until recently it was thought that the increase in electrical resistivity resulting from plastic deformation of polycrystalline copper does not recover to a measurable extent prior to recrystallization. This type of behaviour was indicated for instance by the work of Smart, Smith and Phillips[34] with high purity copper wire cold drawn nominally at room temperature (but due to the heat evolved in the die, actually probably at about 60° to 100°C). However, Molenaar and Aarts[35] discovered recently that, if the plastic deformation is carried out at a low temperature, the increase in resistivity is larger and that it recovers to an easily measurable degree at temperatures at which no recrystallization, and in fact no measurable change in the mechanical properties is taking place. The increase of the effect of plastic deformation on the resistivity with decreasing temperature of deformation has been recently demonstrated more fully for several metals[36]; the effect for Ni is shown in fig. 7. Whether the total amount of this increase is recoverable, or the non-recoverable part of the resistivity increment also becomes larger with decreasing deformation temperature, is not as yet known. R. R. Eggleston[37] found (fig. 8 (*a*)) that for copper deformed at 4°K the electrical resistivity undergoes considerable recovery when annealed for 1 min periods at temperatures as low as −150°C to 0°C and that when the temperature rises to about 100°C this recovery (in view of the kinetics and the low temperatures and short annealing periods involved, it is extremely unlikely that interface migration processes would play a role here) amounts to approximately 50% of the total resistivity increment originally produced by the deformation (fig. 8 (*a*)). Presumably, in his experiments

the remaining non-recoverable portion of the resistivity increment could have been eliminated by recrystallization on annealing at still higher temperatures, as in the work of Smart, Smith and Phillips, if such annealing treatments had been carried out. In analogy to the effect on the recoverable component of work hardening in aluminium, apparently plastic deformation at increasingly higher temperatures also has a decreasing effect on the recoverable component of the resistivity increment until, as mentioned above, none of the resistivity increment is recoverable prior to recrystallization in the case of pure Cu deformed at about 60° to 100°c.

The results of Molenaar and Aarts[35] indicate that the rate of recovery of the electrical resistivity is greater than that of the flow stress, suggesting that these two properties may be affected by different types of imper-

Fig. 7



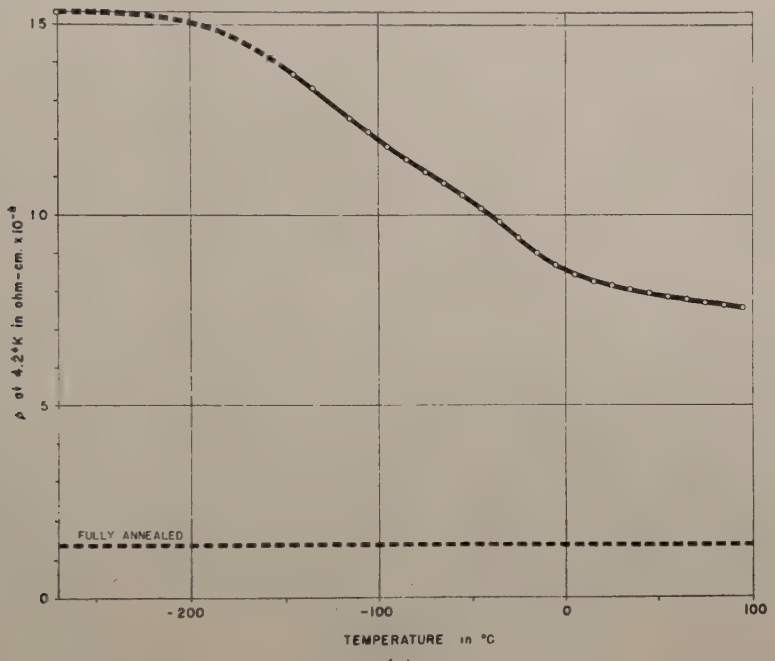
Increase in electrical resistivity in Ni as a function of true strain at 100°c (empty circles), 0°c (crosses), −78·5°c (filled circles) and −183°c (squares). [Ref. (36).]

fections resulting from cold work. Nevertheless, the low temperature drop of electrical resistivity shows kinetics[37] (fig. 8 (*b*)) similar to that of the recovery of flow stress, and of line broadening and of the low temperature release of the stored energy of cold work. Furthermore, all these property changes take place under conditions of annealing time and temperature where no interface migration is observable by microscopic or x-ray diffraction methods, so that they all may be ascribed to recovery. As has been already observed by Tammann[38], the recovery of different properties may take place at quite different rates, so that a cold worked metal may display a whole spectrum of recovery phenomena.* If these correspond to

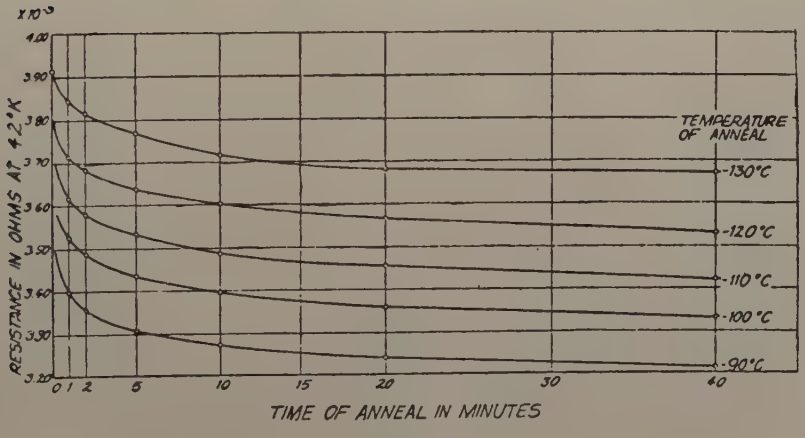
* This is apparently true, even though in some of the early investigations the observed differences in rate may have been due to differences in mechanism, since no clear separation was achieved between recovery and recrystallization.[17]

different elementary processes, it may be that further research will reveal not only variations in the quantitative details of the kinetics, but also differences in the activation energies. On the other hand, if the various recovery phenomena correspond to progressive stages of a single process, their activation energies should be the same.

Fig. 8



(*a*)



(*b*)

Recovery of electrical resistivity as a function (*a*) of annealing temperature, and (*b*) of annealing time at various temperatures, after deformation by twisting at 4·2°K. [Ref. (37) and unpublished work by R. R. Eggleston.]

The main results on recovery to date may be summarized as follows :

(*a*) In general, fractions of certain property changes resulting from plastic deformation, for instance of work hardening, may be recovered in annealing periods and at temperatures where no interface migration can be detected by methods now available. In the case of single crystals deformed in pure shear, without bending, recovery may be carried to completion and the absence of any interface migration or reorientation in the recovered condition is particularly clearly shown by the accuracy with which the original mechanical properties can be re-established.

(*b*) The recoverable fraction of work hardening, of electrical resistivity increment, and presumably also of other property changes, increases with decreasing deformation temperature. The amount of x-ray line broadening, all of which is apparently fundamentally recoverable, also increases with decreasing deformation temperature.

(*c*) The rate of isothermal recovery of each property change is highest at the start, and it gradually decreases as recovery progresses. At least within a certain range, the amount of recovery of the flow stress and of the stored energy of cold work is a linear function of the logarithm of the annealing time, so that, within this range, the rate of recovery is inversely proportional to the annealing time.

(*d*) The kinetics of the recovery of various property changes appears to be very similar, as far as the available data indicate. However, the rate of recovery may be quite different for each property. Thus, apparently the recovery of electrical resistivity is faster than that of x-ray line broadening, and this in turn is faster than the recovery of hardness.

§ 3. Subgrain Growth

3.1. *Subgrains and Subboundaries*

Microscopic evidence for the existence of subboundaries or veining in ferrite deformed at high temperature has been reported by Rawdon and Berglund[39] as far back as 1927–28. Greninger[40] concluded from x-ray diffraction evidence (back reflection Laue patterns) that veining corresponds to a subdivision of the ferrite grains into individual crystal blocks slightly disoriented relative to each other. That the microscopic subboundaries in effect correspond to the limits of disoriented domains in aluminium was demonstrated more recently by Guinier and Lacombe[41]. Collins and Mathewson[42] discovered the appearance of slightly disoriented crystal blocks upon annealing at 500°C of Al crystals previously extended at room temperature ; the Laue asterism streaks, which were continuous after deformation, split into a multitude of more or less well defined dots upon annealing. The information on subgrains up to about 1950 was reviewed by Guinier[8].

3.2. *Subboundary Migration*

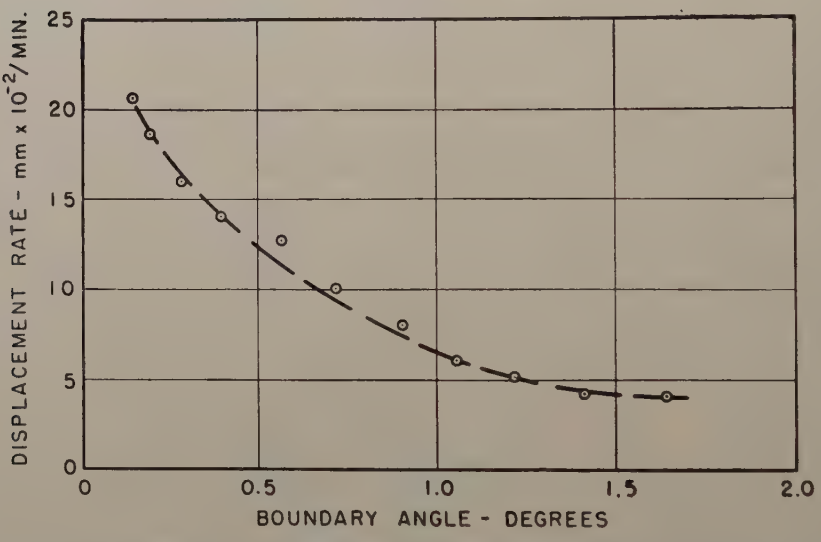
Crussard's early results[43] already indicated that the dots replacing the original asterism streaks in extended Al crystals on annealing at about

350° to 400°C become more sharply defined and less numerous on continued annealing at higher temperatures, particularly above 600°C, suggesting a decrease in the number of crystal blocks, i.e. an increase in the average subgrain size. Also he was the first to note that in Al crystals deformed in tension the subgrain growth on annealing is associated with softening, i.e., a decrease in yield stress. This annealing phenomenon was designated by Crussard as ' recrystallization *in situ* ', meaning *softening without reorientation*. However, in the absence of reorientation, that is the migration of high angle grain boundaries, it is perhaps better not to use the term ' recrystallization '. For this reason, in the present paper the term ' subgrain growth ' has been adopted to describe this stage of the annealing process. That subgrain growth is apparently not always associated with appreciable softening is suggested by the observations of Honeycombe[18]. The movements of subboundaries in aluminium on annealing at around 600°C were studied metallographically by Lacombe and co-workers[44, 45], and by means of a high resolution focusing Laue method by Guinier and Tennevin[46]. The latter investigators found that at the annealing temperatures used the subgrains tend to grow, and that this process is much faster in high purity aluminium than in commercial metal. The work of Dunn and Daniels[47] provided convincing metallographic and X-ray evidence for subgrain growth in plastically bent silicon–iron crystals on annealing in the temperature range of 950° to 1300°C. (figs. 9 and 10, Plates 25 and 26). These workers showed that the process of subgrain growth (increase of the average subgrain size with annealing time and temperature) leads to a decrease in the total subboundary surface energy per unit volume. They pointed out that, as in ordinary grain growth, interface energy may be considered as the driving energy of subgrain growth.

Whether or not the mobility of all very low angle boundaries is in fact particularly high in comparison with grain boundaries of intermediate disorientation, as suggested by Dunn and Daniels, could be definitely decided only by experiments in which the rate of migration of various boundaries is observed in specimens of the same composition, at the same temperature and under the influence of the same driving force. There certainly appears to be a significant difference between subgrain growth and ordinary grain growth at least in regard to the effect of an external stress applied during annealing. Wood and Scrutton[48] discovered that the rate of subgrain growth is greatly accelerated by an externally applied stress, sufficient to produce a small creep strain during the annealing. This effect was later confirmed by Hino, Shewmon and Beck[49] by means of both X-ray diffraction and a microscopic method. If the tension yield stress is accepted as an indicator of subgrain size, as discussed further below, the data of Sherby, Goldberg and Dorn[50] show an example of acceleration of subgrain growth by a factor of over 100 when a creep strain of 0·004 is applied during the annealing of cold worked pure Al at 530°K. On the other hand, Sperry and Beck[51] were unable to detect an

analogous effect of simultaneous stressing on the rate of ordinary grain growth in high purity aluminium. This point is significant since it proves that the often irregularly shaped subboundaries formed in polycrystalline metals on plastic deformation (fig. 11, Plate 27) are in their behaviour more closely related to the crystallographically well defined individual kink boundaries in zinc single crystals, which Parker and his associates[52, 53] have shown to be mobile under external stress (fig. 12, Plate 28) than to ordinary grain boundaries, which are not[54]. The difference in the behaviour of subboundaries and of ordinary grain boundaries is probably not only due to the difference in the extent of disorientation, but more significantly also to the *type* of disorientation[57], as described in the following. Ideal subboundaries of the type described by Burgers[56],

Fig. 13



Mobility of dislocation type boundaries in pure Zn single crystals as a function of disorientation at the boundary. (Rate of migration at 350°c under a constant shear stress of 9·19 p.s.i.) [Ref. (58).]

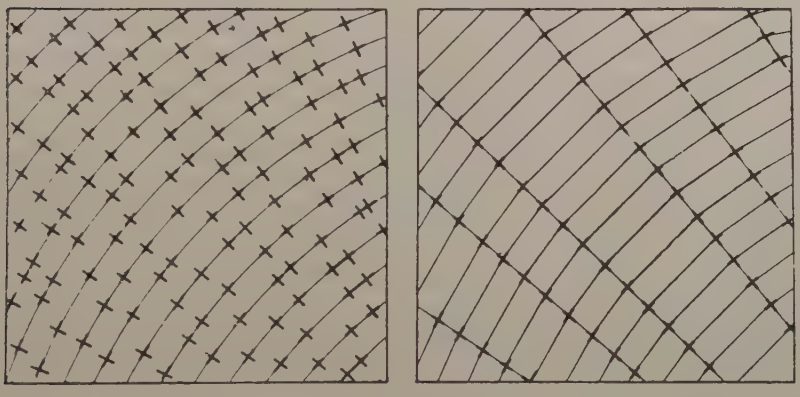
may be constructed of a single set of parallel edge dislocations and the orientation relationship between the two adjoining lattices corresponds to a rotation around an axis parallel to the rational crystallographic direction of the edge dislocations. Such a boundary is highly mobile under stress, even at low temperatures[53, 57], and it may be expected to show relatively high mobility also under the influence of interfacial tension, without externally applied stress. Actual measurement of the mobility of boundaries of this simple type as a function of disorientation was recently carried out successfully by Bainbridge, Li, Edwards, Washburn and Parker[58]. As seen in fig. 13, the mobility gradually decreases with increasing disorientation and shows no sign of reversal of this trend as far

as the measurements were extended, namely up to about 1·6°. On the other hand, when the orientation relationship of two adjoining lattices is of a more general nature, even though the disorientation may be still small, at least two nonparallel sets of edge dislocations are required to form the boundary. In this case, movement of the boundary requires atomic readjustments that can be achieved only by diffusion at the interface[59], even if impurities that pin down the dislocations are disregarded. Consequently, the rate of boundary migration is relatively low, particularly at small disorientations, where the boundary atoms are tightly held in their positions. (For a discussion of the orientation dependence of grain boundary mobility see § 4.5 (*b*).)

3.3. *Mechanism of Subgrain Formation*

As a result of the investigations of Wood and his co-workers[48, 60, 61, 62] it appears to be generally accepted that in creep subgrains form directly

Fig. 14



Distribution of dislocations in a bent crystal (*a*) before polygonization, (*b*) after polygonization, according to Orowan and Cahn. [Ref. (68).]

during the slow deformation at high temperature, and not as a result of annealing subsequent to it. On the other hand, considerable discussion has taken place in recent years concerning the origin of subgrains in metals deformed at low temperatures and then annealed. In view of the work of Cahn[63, 64], who investigated the subgrains formed in cold bent and annealed single crystals of Zn and Al, Orowan and Cahn concluded that the subboundaries are formed during annealing by means of the thermally activated migration of individual dislocations, which were left behind in the plastically bent crystal. On this picture, the continuously bent crystal lattice undergoes sectionwise straightening during annealing, as the

randomly dispersed excess dislocations of one sign rearrange themselves into dislocation walls, or subboundaries, where the orientation change becomes concentrated (fig. 14). A theoretical treatment of this annealing process, named by Orowan ' polygonization ', has been given by Cottrell[65] who showed on the basis of the dislocation theory that the arrangement of edge dislocations one underneath the other, as in a dislocation wall, corresponds to increased stability. In view of the success of these investigators in describing the geometry of subboundaries in cylindrically bent single crystals in terms of dislocation walls, the polygonization mechanism has been quite generally accepted for the formation of subgrains in metals deformed at low temperatures and then annealed. From this point of view, it seemed reasonable for Greenough and Smith[66], for instance, to reinterpret even the subgrain formation in creep as a two-stage process, where the intermediate (continuously bent) condition is not detectable only because of the occurrence of thermally activated ' polygonization ' concurrently with the creep process itself.

However, Wood[67] pointed out that there was no compelling empirical reason for such an interpretation, and he expressed the view[48] that, on the contrary, subgrains may form directly during plastic deformation even at low deformation temperatures. Both Guinier[8] and Cahn[68] noted that, by means of the focusing Laue method of superior angular resolution introduced by Guinier and Tennevin, the latter investigators were able to detect subgrains at a lower annealing temperature (from 450°c up in the case of pure Al) than with the conventional Laue technique used by Cahn (near 600°c). It may be concluded that in all probability the subgrains were already present at an earlier stage of the annealing process, where Cahn's Laue asterism streaks still appeared continuous because of inadequate resolution. Guinier and Tennevin also pointed out[69] that the focusing Laue method's limit of resolution is about 1,3 mm, and that subgrains smaller than that would give continuous Laue streaks rather than resolved individual dots in the x-ray patterns. In fact, further improvement in resolution, achieved by Lambot[70] through the use of a convergent monochromatic x-ray beam focused at a small area of the specimen surface, lead to the detection of a recognizable substructure in extended Al crystals at a still lower annealing temperature (200°c) and even wtihout any annealing above room temperature (fig. 15, Plate 28). As a matter of fact, even in ordinary back reflection Laue patterns, using a fine beam, Honeycombe[18] observed the breakdown of the asterisms into several intensity maxima for Al crystals immediately after 2–6% elongation at room temperature. The presence of a substructure with quite sharp domain boundaries in pure Al deformed at room temperature and not annealed was furthermore convincingly demonstrated by Heidenreich[71] using an electron transmission microscopic method (fig. 16, Plate 29). Kellar, Hirsch and Thorp[72] found that back reflection Debye rings taken of cold worked Al with a microbeam x-ray diffraction technique can be resolved into individual dots. Both of these investigations, and

also the polarized ligth microscopic observations of Beck and Hu[73] gave an approximate subgrain size of 2 microns for pure Al severely deformed at room temperature. No indication of substantial subgrain growth was found[71, 70] on annealing up to 200°C. Heidenreich found that substructure was also present in Al deformed at 78°K and examined at room temperature, although it was distinctly finer than that in the metal deformed at room temperature, namely about 1 micron. More recent microbeam x-ray diffraction work by Gay and Kelly[74] succeeded in detecting the presence of a substructure in polycrystalline Cu, Ni and Fe deformed at room temperature and unannealed (fig. 17, Plate 29). Chen and Maddin[75] found that even a metal with as high a melting point as Mo, shows detectable substructure when a single crystal is deformed at room temperature. Conard, Averbach and Cohen[76] were able to show by means of x-ray diffraction with the specimen held at 78°K, that Zn single crystals deformed at 78°K and not allowed to heat up above that temperature, have a distinct substructure. These observations clearly suggest that substructures are formed on plastic deformation under conditions where thermally activated processes can not play a significant role.

As mentioned above, the interpretation of subboundaries as ' dislocation walls ' or arrays of parallel edge dislocations by Orowan, Cahn[62] and Cottrell[65] was originally associated with the view that such boundaries are formed an annealing by ' polygonization ', i.e., by means of the sectionwise straightening of the crystal lattice continuously curved in plastic deformation. Polygonization was viewed as the thermally activated migration of individual dislocations from their random positions in the deformed crystals and the collection of these dislocations along planes nearly perpendicular to the active slip direction. In view of the results described in the previous paragraph, it now appears quite certain that dislocation walls are actually formed during the deformation process itself, even at very low deformation temperatures. It is, of course, not necessary to assume that *all* the excess dislocations of one sign are necessarily collected into dislocation walls during the deformation. The lack of sharp extinction contour lines[71] and the width of the individual x-ray micro beam diffraction spots[77] suggest that some plastic strain may be present within the subgrains themselves in the cold worked condition. Nevertheless, the fact, shown by the various experiments mentioned above, that metal crystals are in effect ' fragmented ' into rather distinct lattice domains of differing orientations directly on plastic deformation, apparently without requiring much or any thermal activation, suggests that a mechanism other than polygonization is involved. Whether or not some additional subboundaries actually do form on subsequent annealing by polygonization can not be decided with certainty at the present time. This question is intimately related to the as yet indefinite interpretation of the ' background ' darkening between the separated spots in x-ray diffraction patterns, such as fig. 17. This background apparently originates from diffraction by crystal regions of orientation intermediate

between those giving rise to the separated diffraction spots. These intermediate orientations may possibly be represented by distinct subgrains too small to be resolved with the diffraction method used, or by continuously bent crystal regions. In the event that the former interpretation should prove correct, the apparent continuity of the background would be eliminated by subgrain growth. If the latter interpretation be true, the disappearance of the continuous background on annealing may result from the movement of excess dislocations of one sign from such bent regions either to subboundaries previously formed in the vicinity during deformation or to subboundaries newly formed on annealing by polygonization.

As to the mechanism of subboundary formation on plastic deformation, it is significant that the deformation mechanism known as ' kinking ' according to Orowan[79], Jillson[80], Hess and Barrett[81], Wsahburn and Parker[52] and Li, Edwards, Washburn and Parker[53], results in the formation of dislocation walls identical with those one expects to form by thermal activation on the basis of the postulated polygonization mechanism. In view of the known mobility of edge dislocations under the effect of shear stress even at low temperatures, and of the stability of their arrangement in the form of dislocation walls according to Cottrell[65], it seems rather reasonable to assume that a large majority of the excess dislocations of one sign will so align themselves during the deformation process itself. Recently Bainbridge, Li, Edwards, Washburn and Parker[58] observed that two parallel kink boundaries in a Zn single crystal, migrating on application of an external shear stress, may join to form a single kink boundary of greater disorientation, even when the crystal is strained at a temperature as low as 78°K, and that a 1° kink boundary may trap additional dislocations, which pass through the crystal on straining, and thereby increase the boundary angle to as high as 15°. On the basis of the various observations described, it appears quite probable that in metals the subboundaries form on plastic deformation essentially by the kinking mechanism—an athermal process.

The formation of a kink boundary requires slip more or less evenly distributed over a large number of parallel slip planes. For a subboundary with a disorientation of $\frac{1}{2}°$ the required average distance of the active slip planes in Al would be of the order of $0{\cdot}02\,\mu$, that is much smaller than the observed distances of ordinary slip bands (order of $1–10\,\mu$). However, this requirement represents no difficulty [96] for the proposed kink mechanism of subboundary formation, since it has been known already from the early work of Masing and Polanyi[82] that single crystals of Sn can be sometimes extended 1000% without the formation of any visible slip bands. That such ' fine slip ' in effect may occur to a substantial degree even in the presence of slip band formation, was shown recently for Al by Kuhlmann-Wilsdorf and Wilsdorf[83]. The above mentioned results of Parker and his associates on the movement of kink boundaries under externally applied stress in single crystals of Zn provide additional proof for the

occurrence of fine slip; the movement of a dislocation wall under such conditions can take place only by that mechanism.

3.4. *Factors Affecting the Subgrain Size*

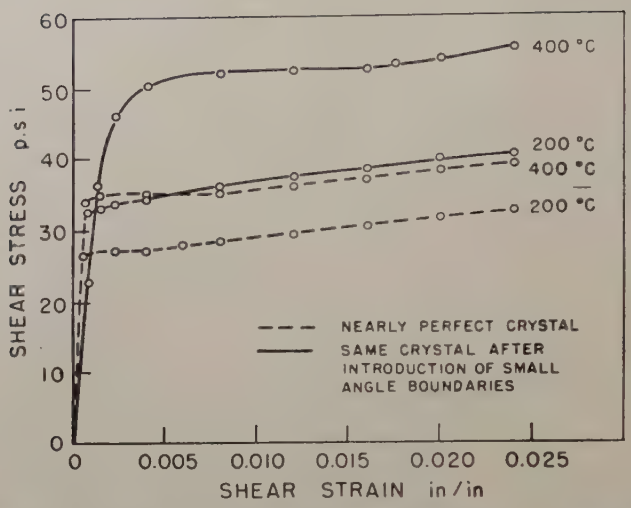
Although the formation of subboundaries is apparently essentially athermal, as discussed, nevertheless there is a fairly clear indication that a thermally activated process is somehow involved at least in determining the subgrain size. Wood and Rachinger[62] have shown that the subgrain size resulting from creep in pure polycrystalline Al increases with increasing deformation temperature from 20° to 350°C and with decreasing strain rate from 10% per min to 0·1% per h. Such an effect of the deformation temperature was also found by Heidenreich[74] between 78° and 300°K, at much higher strain rates than those used in the experiments of Wood. The information available concerning this effect is not sufficiently quantitative to allow detailed interpretation at the present time. In any event, the phenomenon of subgrain growth, discussed above, definitely depends on thermal activation; the results of Dunn and Daniels[47] are particularly clear cut in showing the increase in subgrain size on annealing both with increasing time at constant temperature and with increasing annealing temperature.

The effect of solute atoms on the subgrain size formed on plastic deformation, and particularly on subgrain growth, are not well known as yet. Heidenreich observed[71] that 4% of Cu dissolved in Al greatly reduces the subgrain size formed on deformation at room temperature. Hazlett and Parker[84] noted that the addition of 1 atomic per cent Ti to Ni refined the substructure under otherwise similar conditions. One may expect that the solute atoms may be also quite effective in preventing the migration of dislocation walls as they are known from the work of Cottrell[85] to immobilize individual dislocations. On this basis, a very great decrease in the rate of subgrain growth would be expected to result from the presence of a relatively small number of solute atoms. There appear to be no data available at the present time to test this expectation.

3.5. *Effect of Substructure on Mechanical Strength*

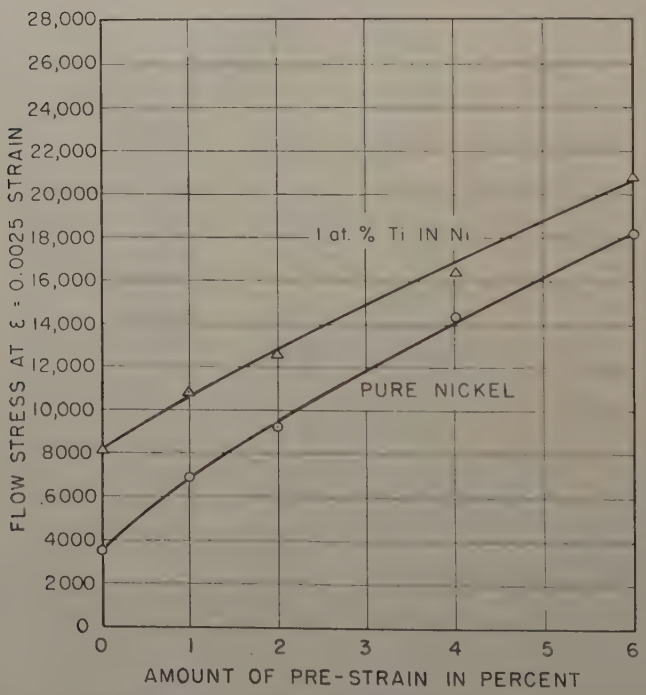
There is now a rather convincing body of evidence[86] indicating that the presence of substructure has a great effect on the mechanical properties. As shown in fig. 18, Li, Washburn and Parker[16] found that in a Zn single crystal the flow stress in shear is greatly increased if a substructure is introduced by a ball indentation at 350°C. Annealing at 400°C does not eliminate this large scale substructure (fig. 19, Plate 30), although other types of imperfections, which cause work-hardening when Zn crystals are sheared without bending, are known to be annealed out readily at much lower temperatures[15, 16]. Since the high shear flow stress was actually observed after a 400°C anneal, it must have been a result of the substructure and not of the other types of imperfections mentioned. In their

Fig. 18



Shear stress–strain curves for Zn single crystals before (– – –) and after (———) introducing substructure by indentation in the [0001] direction at 350°c. Specimens were annealed at temperatures indicated and cooled to room temperature in 15 minutes, before testing. [Ref. (16).]

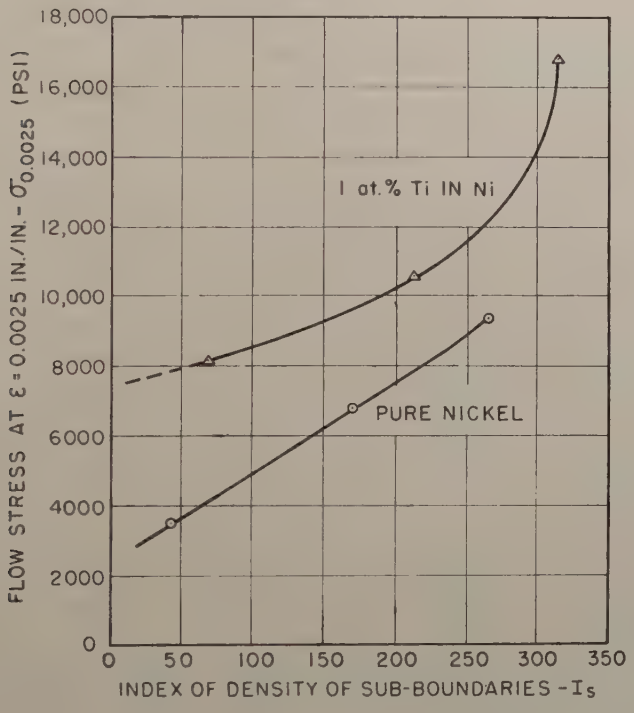
Fig. 20



Flow stress at 0·0025 strain for pure Ni and for a Ni–Ti alloy as a function of prestrain. Specimens were prestrained at room temperature and annealed for 1 hour at 800°c before testing. [Ref. (84).]

recent work with Ni and with Ni–Ti solid solution alloys, Hazlett and Parker[84] found that a few per cents of prestraining in tension followed by annealing at 800°C (which is presumably sufficient to recover work hardening due to imperfections other than subgrains) greatly increases the yield stress in a subsequent tension test. For instance, prestraining to 5% raises the yield stress for pure Ni from about 3500 p.s.i. to 18 000 p.s.i. (fig. 20), apparently as a result of decreasing subgrain size. Figure 21 shows the variation of yield stress as a function of the ' index of amount of substructure ' for the same specimens. Sherby, Goldberg and Dorn[50]

Fig. 21

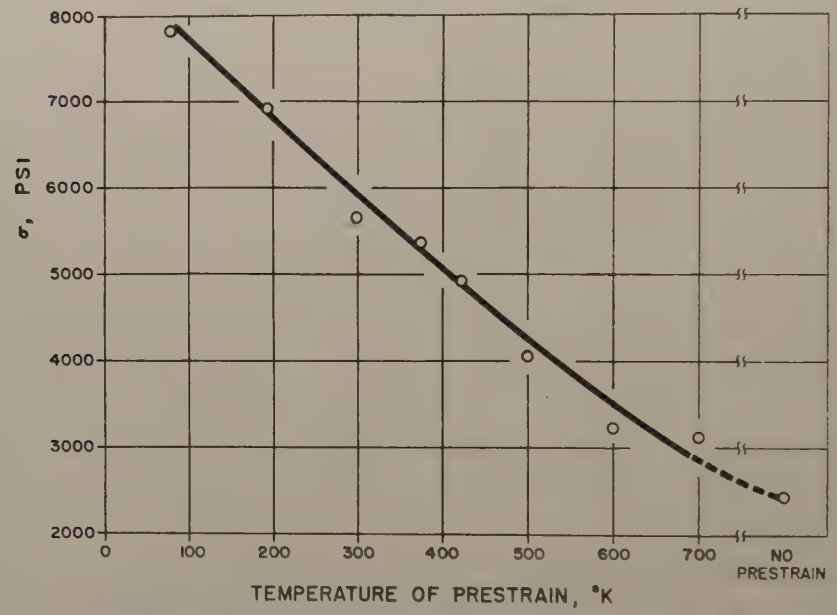


Same flow stress data as in fig. 20, plotted as a function of the density of subboundaries resulting from prestraining to various amounts and annealing for 1 hour at 800°C. [Ref. (84).]

obtained data for pure Al, which may be also interpreted as an illustration of the effect of subgrain size on the nonrecoverable component of work hardening. Tension specimens were prestrained under varying conditions, so that substructures of varying fineness were presumably obtained. The recoverable component of work hardening was eliminated either by carrying out the prestraining at temperatures above 422°K, or by annealing

for 2 hours at 530°K, after prestraining at lower temperatures. It may be assumed that the remaining work hardening in these specimens was practically all due to the subboundaries. The observed increase of yield stress in tension tests at 78°K from about 1300 p.s.i. to 7400 p.s.i., for instance with decreasing temperature of prestraining (fig. 22), may be ascribed to the progressive decrease in subgrain size. In other experiments, specimens prestrained at 78°K and presumably possessing a very fine substructure were annealed under varying conditions. In view of the effect of such conditions on the substructure, as known qualitatively from the previous investigations of Wood and his associates, the resulting changes in mechanical properties may be again interpreted as consequences

Fig. 22



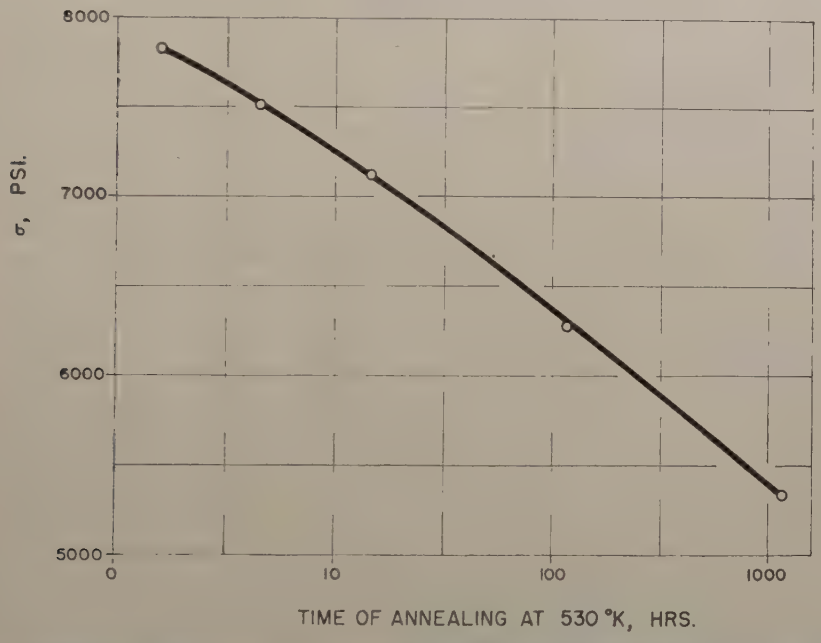
Flow stress for pure Al at 0·01 strain at 298°K after prestraining 15% at temperatures shown and annealing for 2 hours at 530°K. The increase in flow stress with decreasing temperature of prestrain is presumably due to the increasing density of subboundaries, since the annealing probably removed all recoverable work hardening. [Replotted from Ref. (50).]

of variations in subgrain size. Thus, fig. 23 shows the decrease in tension yield stress as a function of the period of annealing at 530°K, presumably as a result of subgrain growth. Although the above interpretation appears to be a very probable one, it clearly lacks definitiveness for two reasons: (1) the absence of adequate subgrain size measurements, and (2) lack of proof of the absence of lattice imperfections other than subboundaries.

It has been observed microscopically by Beck and Hu[73] that subgrain growth occurs to a significant extent in the annealing of polycrystalline

high purity Al, and that it serves as a mechanism of annealing by which the orientation of the deformed material may be partially retained. Recently, some instances became known where complete softening in highly deformed polycrystalline metals occurs essentially without reorientation. Thus, Semchyshen and Timmons reported[87] that in rolled Mo sheet, the texture of which has substantially a single component of orientation, this orientation is retained without appreciable change on soft annealing, although the material then consists of newly formed, microscopically well resolved grains. Since the retainment of a single component texture on annealing excludes ordinary high angle grain boundary migration,

Fig. 23

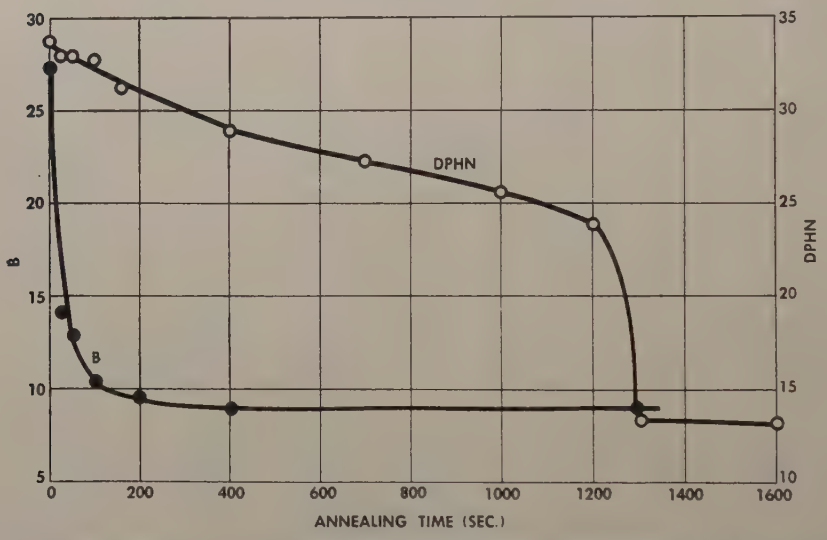


Flow stress for pure Al at 0·01 strain at 298°K after prestraining 15% at 78°K and annealing at 530°K for periods indicated. The decrease in flow stress with continued annealing is presumably due to decreasing density of subboundaries attained by subgrain growth during annealing. [Replotted from Ref. (50).]

Semchyshen and Timmons' results may be best interpreted as softening by means of low angle boundary migration, i.e., subgrain growth. It can be expected that, in the course of this process, the disorientation between neighbouring subgrains gradually becomes larger, and the character of the interfaces more and more approaches that of ordinary grain boundaries in a highly oriented structure, as for instance in Cu with cube texture. In such cases subgrain growth may change over into ordinary grain growth quite gradually and it may be rather difficult to separate these processes.

Another instance apparently similar to this was observed by McHargue[88] with rolled Ti, where again complete softening occurred without change in orientation. (In this case, the identification of the process as subgrain growth is somewhat less definite, since the texture consists of two components.) McGeary and Lustman[89] obtained data that may be interpreted as indicative of 50% loss of work hardening in cold-rolled polycrystalline Zr on annealing at 400°C, essentially without reorientation. Although all these results appear highly significant, the definition of orientation in these polycrystalline textures is not very sharp and, therefore, the detection of reorientation is not very sensitive. An instance of complete softening without reorientation in a deformed single crystal of silicon–iron was mentioned by Dunn[90].

Fig. 24



Vickers hardness (D.P.H.N.) and residual x-ray line broadening (B) for pure Al single crystal rolled 80% on (110) plane in [1$\bar{1}$2] direction and annealed at 350°C for periods indicated. [Ref. (30).]

In a recent paper, Lutts and Beck[30] presented data for pure Al single crystals rolled 80% on a (110) plane and in a [112] direction, where a very sharp deformation texture results with the same orientation as that of the original crystal[91]. On isothermal annealing at 350°C it was found that the line broadening recovers completely at an early stage of the annealing, when approximately 75% of the work hardening is still present (fig. 24). This indicates that at least the considerable remaining fraction of the work hardening is not associated with other types of lattice imperfections detectable by line broadening measurements (such as strain or stacking faults) and should be presumably attributed to the presence of subboundaries. Furthermore, it was found that after annealing to the

completely soft condition at 350°c the material has exactly the same very sharp single orientation texture as before annealing, so that, in this case, softening is definitely not associated with reorientation, that is the migration of ordinary high angle grain boundaries (fig. 25, Plate 31). The elimination without recrystallization of work hardening presumably due to subboundaries should be attributable to subgrain growth, although on this crucial point direct evidence is still missing.

The principal conclusions with regard to subgrains on the basis of results to date may be summarized as follows :

(*a*) Subgrains, that is crystal domains slightly disoriented with regard to each other, are formed in metallic crystals on plastic deformation, presumably mainly by kinking, even at temperatures as low as 1/10 of the absolute melting point. The only deformation so far known not to be connected with subgrain formation is pure shear (without bending) in single crystals.

(*b*) The subgrain size formed under otherwise similar conditions decreases with decreasing deformation temperature, increasing strain rate, increasing total strain (within certain limits) and with the addition of solute atoms.

(*c*) Annealing under suitable conditions may produce subgrain growth, leading to a decrease of total subboundary energy per unit volume, and thus to increased thermodynamic stability. Subgrain growth is often prevented, particularly in deformed polycrystalline metals, by the premature onset of recrystallization, which wipes out the substructure.

(*d*) The rate of subgrain growth is greatly increased by an external stress applied during annealing, sufficient to produce a slight strain. An analogous effect in ordinary grain growth was not detected.

(*e*) Large work hardening effects are connected with the substructures produced in plastic deformation. The fraction of the total work hardening due to this effect increases with the deformation temperature for any one metal. Subgrain growth on annealing may account for at least a part of the relief of work hardening. The extent of relief by this mechanism depends mainly on its rate, in relation to that of recrystallization. Some instances of complete softening without recrystallization are now known, for which subgrain growth is presumably responsible.

§ 4. Recrystallization

4.1. *Primary Recrystallization, Gradual Grain Growth and Coarsening*

The classical picture of recrystallization, as the term has been quite generally used in the English language literature[92], involves the growth of new, strain-free grains at the expense of the cold worked matrix. On this picture, the following changes take place simultaneously, as a result of the growth of the new grains : (*a*) *release of stored energy of cold work* in proportion to the volume of cold worked matrix replaced by new, unstrained grains ; (*b*) *softening*, as work-hardened material is replaced by soft annealed

grains. In the partially recrystallized material hard and soft areas coexist side by side, as may be ascertained by means of microhardness measurements[94] ; (*c*) *local reorientation*, as an ordinary high angle grain boundary migrates and a deformed grain of a certain orientation is replaced by an annealed grain of a different orientation (fig. 26, Plate 30).

However, recent investigations made it very clear that these changes do not necessarily occur simultaneously[95, 86]. As discussed in § 3.5, instances are now known, where complete softening takes place prior to reorientation, that is essentially without the migration of high angle grain boundaries ; recrystallization may then subsequently occur without further softening. There have been also reported recently some instances where a large part, or substantially all, of the stored energy of cold work is released prior to softening and to reorientation, as will be discussed in further detail below. In view of the variable and sometimes negligible, amount of residual work hardening and strain energy left in the matrix when it is being absorbed by the new grains, the phenomena of recrystallization in the above defined narrower sense (also called ' primary recrystallization '), can not be sharply separated from the types of boundary migration (' secondary recrystallization ' or ' grain growth ') where work hardening or strain energy admittedly play no significant role. Because of the lack of sharp distinction between these processes, it is perhaps best to adopt the nomenclature used by Alterthum[97], van Arkel[98] and Burgers[6], where all boundary migration phenomena connected with the movement of ordinary high angle grain boundaries are designated by the generic term recrystallization. In the following, the term recrystallization will be used in this general sense. It has been customary to distinguish, largely on the basis of kinetics, between two types of grain growth phenomena. In ' gradual grain growth ' or ' normal grain growth ' many grains in a soft annealed matrix grow at the expense of others, resulting in an increase of the average grain size, without any substantial change in the grain size distribution. On the other hand, in ' abnormal grain growth ', ' secondary recrystallization ', ' discontinuous grain growth ' or ' coarsening ' relatively few grains which become very large, absorb a soft annealed matrix, which itself remains fine grained until absorbed. Characteristic of this process is the marked grain size contrast between the coarse grains and the matrix (' duplex structure '). In the following, the terms ' primary recrystallization ', ' gradual grain growth ' and ' coarsening ' will be used to designate the three rather roughly outlined types of recrystallization phenomena mentioned above.

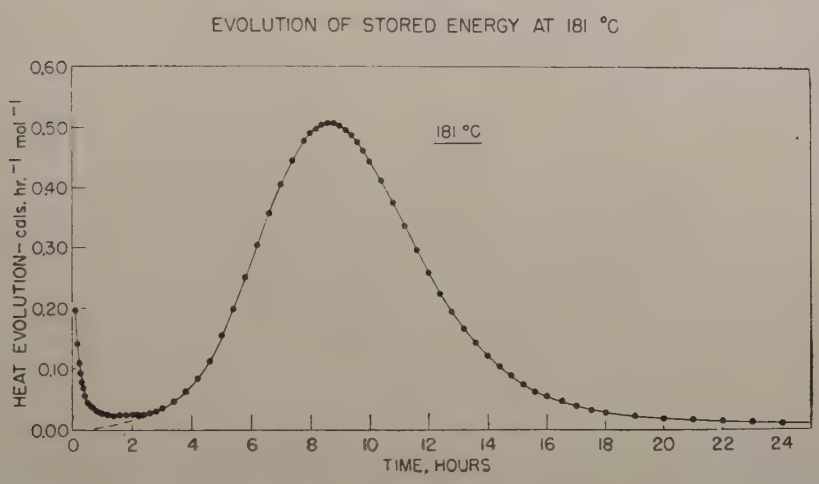
4.2. *Driving Energy and Energy Release*

(*a*) *Primary Recrystallization*

As stated in the Introduction, the energy absorbed in a metal specimen during cold working is of the order of 1–10% of the energy expended to bring about plastic deformation[1, 2, 3, 4, 5]. It has been quite generally held[99] that most or all of this stored energy of cold work is released during

primary recrystallization, and that it represents the driving energy for this process. However, the recent investigation of Borelius and his associates[26] showed that at least a portion of the stored energy may be released at relatively low temperatures and short annealing times, where recrystallization does not occur. Suzuki reported[2] that in 99·96% pure copper deformed in compression to various amounts and heated at the rate of 2°c per minute the energy release is practically completed at a temperature (270°c), where the softening, as measured by indentation hardness or compression yield stress, just begins. It may be assumed that in this case the energy release is associated mainly with processes occurring prior to grain boundary migration. Preliminary results by Stansbury[100] with a silver bearing copper appear to confirm Suzuki's

Fig. 27



Isothermal release of stored energy of cold work at 181°c from pure Cu extended 15%. (Unpublished data by P. Gordon.)

findings. On the other hand, Clarebrough, Hargreaves, Michell and West[3] found with phosphorized coppers of various purity that on heating up at 2° per min, by far the largest part of the energy release took place concurrently with softening and with microscopically observable grain boundary migration. In an as yet unpublished investigation, Gordon[5] observed that in high purity copper deformed in tension the largest part of the energy release was also associated with recrystallization. Gordon's isothermal energy release curves at 200°c (figs. 27) are interesting in that they show an initial energy release with typically recovery type kinetics, similar to that found by Borelius and his associates[26], followed by a clearly separated large peak of energy release corresponding to recrystallization, and with kinetics typical of the latter process. The cause of the

discrepancy between the various investigators' results has not as yet been clarified. The conditions of deformation, the impurities in the copper, and the method of measurement were different in each case.

In contrast to their results with copper, Clarebrough, Hargreaves and West[101] find that deformed nickel releases about 60% of the stored energy of cold work prior to the onset of recrystallization. This suggests that the unexplained differences in the results for copper, mentioned above, may be due to real differences in the behaviour of the various specimens rather than to the methods of measurement. At the present stage of this work it appears safe to conclude only that in some cases most of the energy release is associated with the absorbtion of the deformed matrix by new grains, as in the classical picture of recrystallization, but that this is not true in other cases. Under certain conditions, new grains may grow in a cold worked matrix, which has already lost the largest part of the stored energy of cold work in the early phases of annealing without grain boundary migration. The remaining minor fraction of the stored energy of cold work, which in such cases serves as the driving energy of the recrystallization process, is probably associated with the network of subboundaries in the matrix, as suggested by C. S. Smith[102]. In the experiments of Lutts and Beck, referred to in the section on Subgrain Growth, ' recrystallization ' i.e. the migration of high angle boundaries takes place at a stage of annealing after the completion of softening (figs. 24 and 25), where not only recovery had been completed, but presumably even some of the subboundaries had been annealed out by means of subgrain growth. It is very likely that in this case, too, most of the stored energy of cold work was released prior to recrystallization.

(*b*) *Grain Growth and Coarsening*

The early suggestion of Ewing and Rosenhein[103], Tammann[104], Vogel[105] and Alterthum[97] that the driving energy of grain growth in soft annealed metals is the interface energy associated with the grain boundaries has received strong support in recent years through the work of Harker and Parker[106], C. S. Smith[107] and J. E. Burke[108]. Burke and Schiau[109] have shown that small residual strains, such as might remain in an annealed metal, have little if any effect on gradual grain growth. As estimated from the value of the interfacial energy per unit interface area and from the change in total interface area per unit volume, the energy released when a metal with an average grain diameter of 10^{-3} cm undergoes grain growth to an average grain diameter ten times larger is of the order of magnitude of 0·2 cal/mole. This is in the order of 100 times smaller than the stored energy of cold work, which is released prior to or during primary recrystallization. The driving energy of grain growth or coarsening is, of course, even smaller if the initial grain size is, as commonly happens, larger than that assumed above. In view of the small magnitude of this energy, it is understandable that no successful attempts to measure it directly are recorded in the literature.

4.3. *Kinetics*

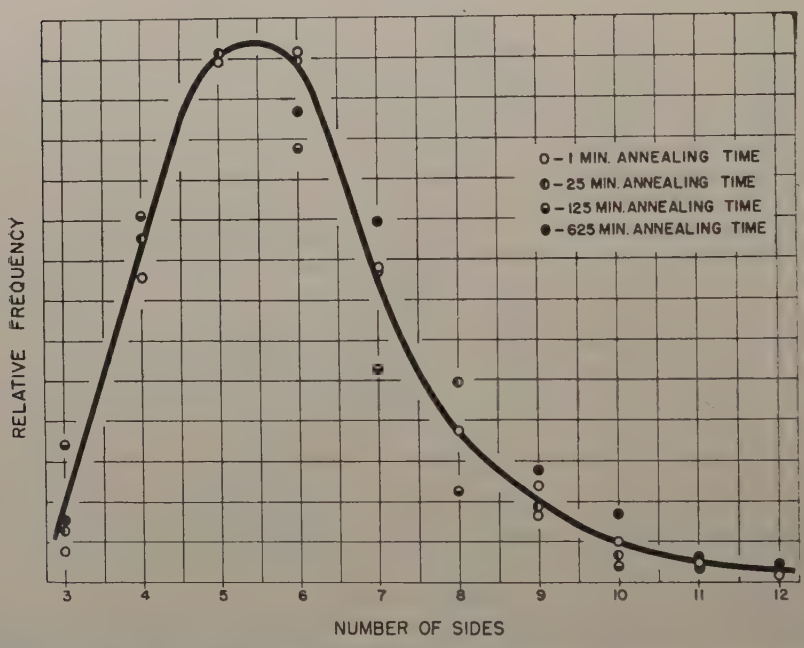
(*a*) *Gradual Grain Growth* (*Normal Grain Growth*)

Vogel[105] showed a tendency of grain boundaries to migrate in such a way as to decrease their curvature, and ascribed this tendency to the requirement of minimizing the interface energy associated with the boundary. Harker and Parker[106] in their more complete analysis later found that in addition to the tendency of grain boundaries to straighten themselves, approach to equilibrium requires also the migration of grain boundaries so as to minimize the resultant surface tension forces exerted at a grain edge by the three grain boundaries intersecting at it. They recognized that grain boundary migration is brought about by annealing whenever the system of grain boundaries in a metal does not satisfy both equilibrium conditions mentioned. Smith[107] showed that true equilibrium arrangement of grain boundaries in three-dimensional space would be realized only if each grain had the shape of a minimum-area tetrakaidecahedron, in complete analogy to the stable arrangement of soap films in a froth. He, and later in more detail Burke[108], pointed out that a deviation from true boundary equilibrium configuration at any point in a polycrystalline metal must result in continued boundary migration in the whole specimen because of the tendency of the deviation to spread to other areas. Since the deviations from equilibrium are not eliminated from the metal despite the boundary migration, true equilibrium is in practice not achieved until the whole specimen is converted into a single crystal. As described by Smith[110], a curved grain boundary will gradually migrate on annealing toward its center of curvature, requiring further continual readjustments at the edges. This process takes place relatively slowly, the rate in a specific case depending on the local curvature of the boundary and on the temperature. At the moment this adjustment process leads to a locally highly unstable configuration where more than three surfaces meet at an edge, or more than four edges meet at a point, the grain boundaries rapidly rearrange themselves, approaching new positions of local equilibrium. As confirmed by direct microscopic observation of grain boundary migration in pure Zn slightly above room temperature, locally the process consists in alternating slow migration and quick readjustment. Small, tetrahedral grains have convex boundaries. These necessarily move toward the center of such a grain, until it disappears. As the number of grains per unit volume decreases the average grain size increases.

In pure metals or solid solutions with a small initial grain size, annealing leads to an increase of the average grain size, but it leaves the type of the frequency distribution of the grain sizes more or less unchanged. Hu and Beck[125] made a quantitative study of this phenomenon for isothermal grain growth in pure aluminium. Planimetric measurement of the area of each of approximately 500–1000 grains in the center section of strip specimens at three successive stages of grain growth gave grain size distribution curves that are similar to each other (fig. 29), and can be made to coincide fairly well by proportional changes in the grain size scale (fig. 30). As seen above, the increase in the average grain size is

accomplished by the disappearance of the smallest tetrahedral grains. At any point during the annealing some larger grains with more than four boundary surfaces are being partially absorbed by their neighbours until they lose some of their boundary surfaces, eventually become tetrahedral, and finally disappear. The largest grains, i.e. those with the largest number of boundary surfaces are the ones that tend to grow most. During grain growth the number of grains of a given size range continually decreases, since some grains leave the group by becoming too small or too large, while at the same time this number is increased the grains just becoming small enough or large enough to fall within

Fig. 28



The relative frequency of grains with various numbers of sides in plane section remains essentially unchanged during grain growth in pure Al. Rolled Al specimens annealed at 500°c for periods shown (see fig. 32 and Ref. (113), micrographic tracings by H. Hu, unpublished data on number of sides by P. R. Sperry and N. Gottesmann).

the size range considered. As seen in fig. 29, the net result is that the frequency of the largest grains at any time increases, and that of the smallest grains decreases. The frequency of grains of certain intermediate sizes may first increase and then decrease in the course of continuing grain growth. Although the local boundary movements are quite discontinuous, the overall change with annealing time of the structure as a whole, i.e. of the statistics of the grain size distribution, is a very gradual or continuous one. The frequency of grains with various numbers of sides (in a plane section through the specimen) remains practically unchanged during gradual grain growth (fig. 28).

The increase of the average grain size with time on isothermal annealing follows a parabolic relationship and it can be represented empirically[111] by the following formula :

$$D^{1/n}-D_0^{1/n}=C\,.\,t, \qquad \ldots\ldots \quad (2)$$

where t is the annealing period for grain growth, D_0 is the average grain diameter before and D is that after annealing, C and n are parameters, which depend on the temperature, but are independent of D. Figures 31 and 32 give experimental data for high purity brass and aluminium. For a given temperature the rate of grain growth depends only on the instantaneous grain size[114] (fig. 33). From (2) the instantaneous rate of grain growth

$$\frac{dD}{dt}=n\,.\,C\,.\,D^{1-1/n}. \qquad \ldots\ldots\ldots \quad (3)$$

The value of n for pure brass is approximately 0·5, independently of the temperature. For pure aluminium n varies from 0·056 at 350°C to 0·322 at 600°C, extrapolating to a value of near 0·5 at the melting point. Since n is always less than one, (3) gives rates of grain growth decreasing with increasing grain size (figs. 34 and 35).

As seen in fig. 36, if log D is plotted against the log of the annealing time for cold rolled strips of brass, the initial points often deviate from a straight line relationship. This is to be expected from (2)[111], since the total isothermal annealing time includes a period for recrystallization and this, in general, differs from the hypothetical annealing time that would be required to produce by grain growth from an initial grain size of zero, that grain size which is in fact produced by primary recrystallization. Depending on the amount of deformation, and thus on the time necessary for recrystallization, this initial deviation may be either upward or downward (fig. 36). However, if the time scale is shifted so as to plot the log of the time for grain growth alone, but include in this also the time for the hypothetical grain growth from zero size to the recrystallized grain size, then in accordance with (2), all points fall on a straight line (fig. 37). The results here reported were obtained by Towers and Beck[112], who carried out the necessary correction in the time scale for each curve by means of independent determinations of the time for complete recrystallization and of the grain size as recrystallized.

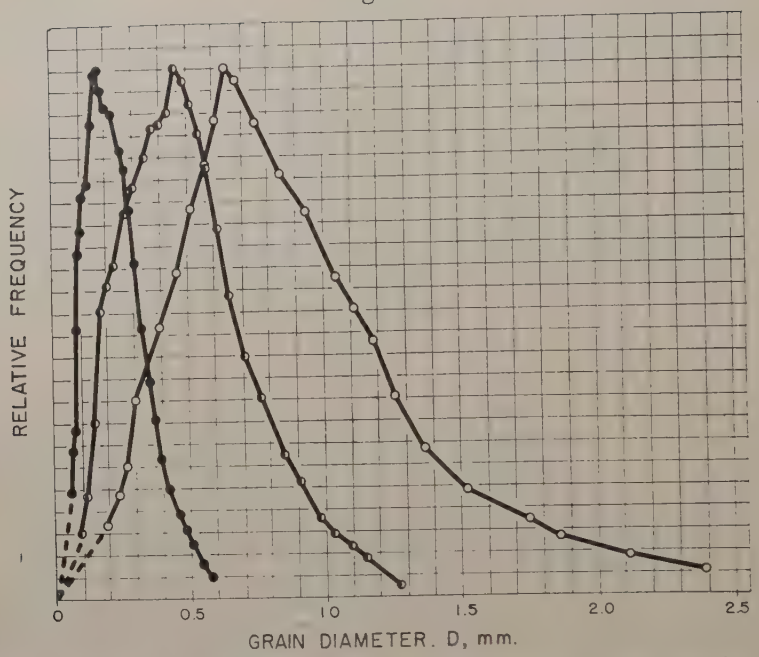
A relationship, such as (2) with $n=0{\cdot}5$, may be derived[113] if one assumes that the instantaneous rate of grain growth is proportional to the instantaneous value of the excess free energy per unit volume associated with the grain boundaries. Since this energy may be taken as proportional to the grain boundary surface area per unit volume, it follows :

$$\frac{dD}{dt}=K.\frac{1}{D} \qquad \ldots\ldots\ldots \quad (4)$$

from which

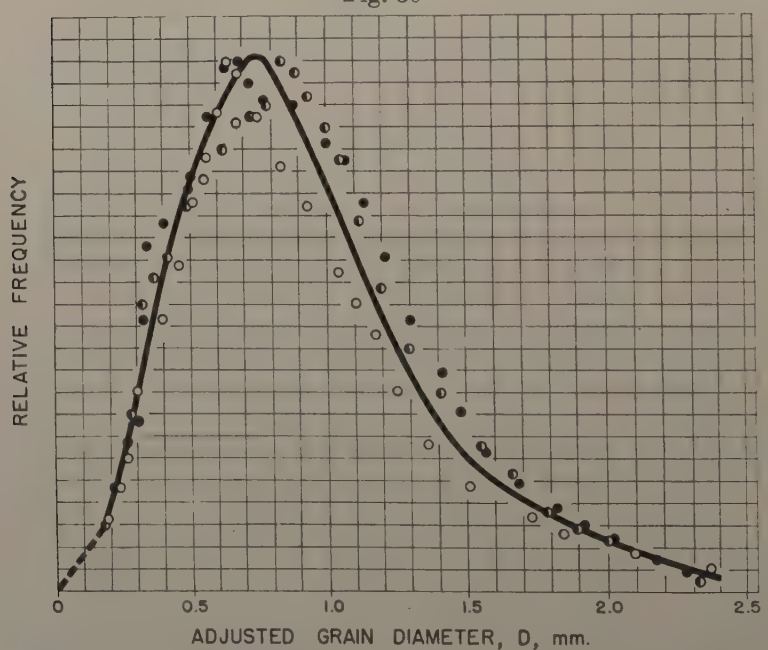
$$D^2-D_0^2=C\,.\,t. \qquad \ldots\ldots\ldots \quad (5)$$

Fig. 29



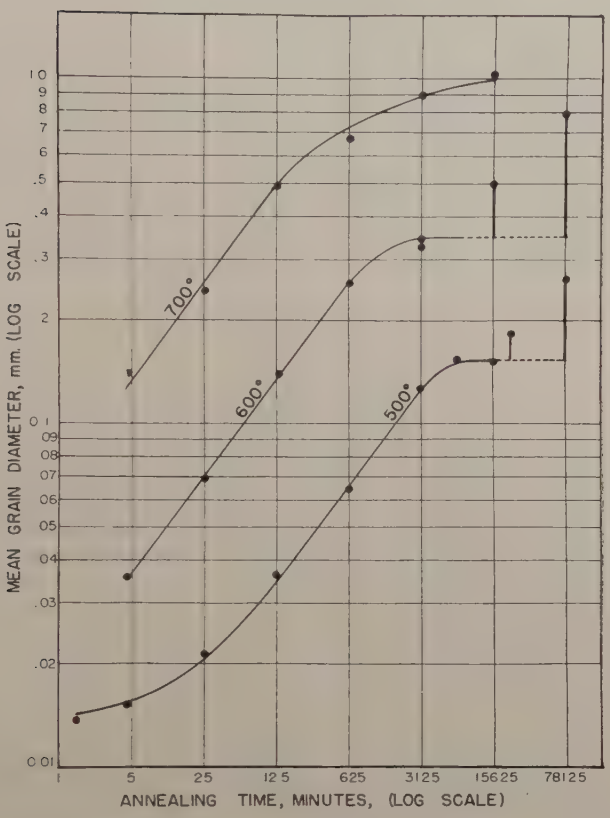
Relative frequency of grains of various diameters in plane section, after grain growth at 500°c for 1 minute (solid circles), 25 minutes (half-filled circles) and 625 minutes (empty circles). Same pure Al specimens as in fig. 28. See also fig. 32. (Unpublished data from 500–1000 grains in each specimen by H. Hu and P. A. Beck.)

Fig. 30



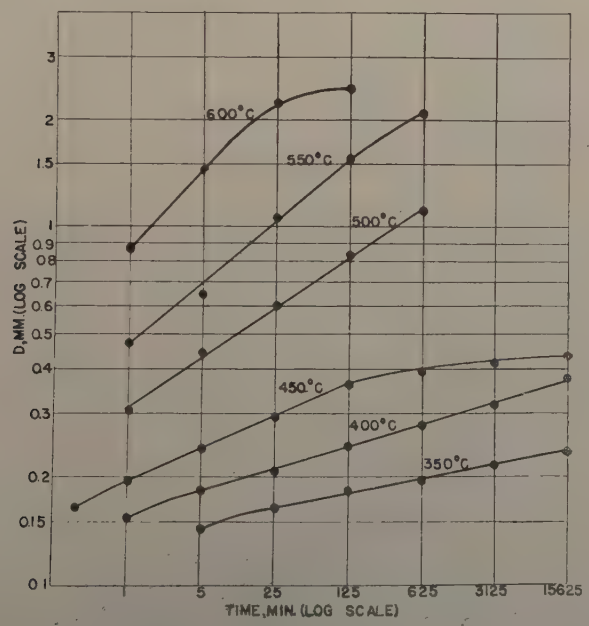
The relative frequency data of fig. 29, replotted over grain diameter adjusted proportionally to make the average grain diameter in each specimen 0·96 mm. The type of grain size frequency distribution remains essentially unchanged during grain growth. (Unpublished data by H. Hu and P. A. Beck.)

Fig. 31



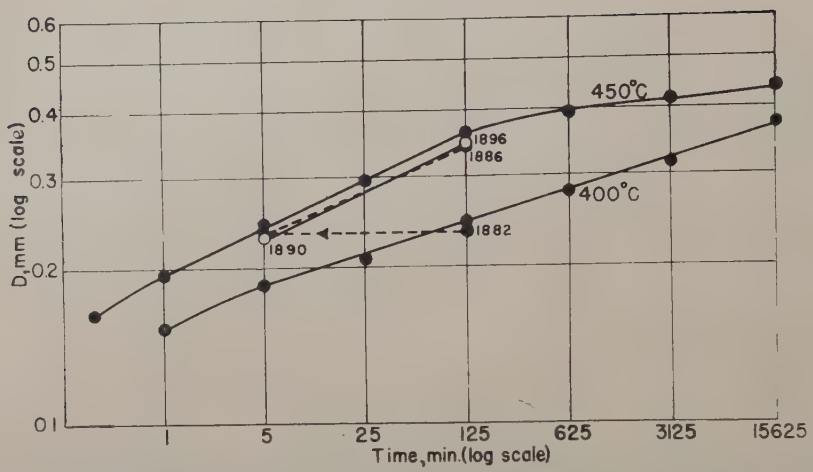
Isothermal grain growth in high purity 70–30 brass. (Unpublished data by J. Towers and P. A. Beck.)

Fig. 32



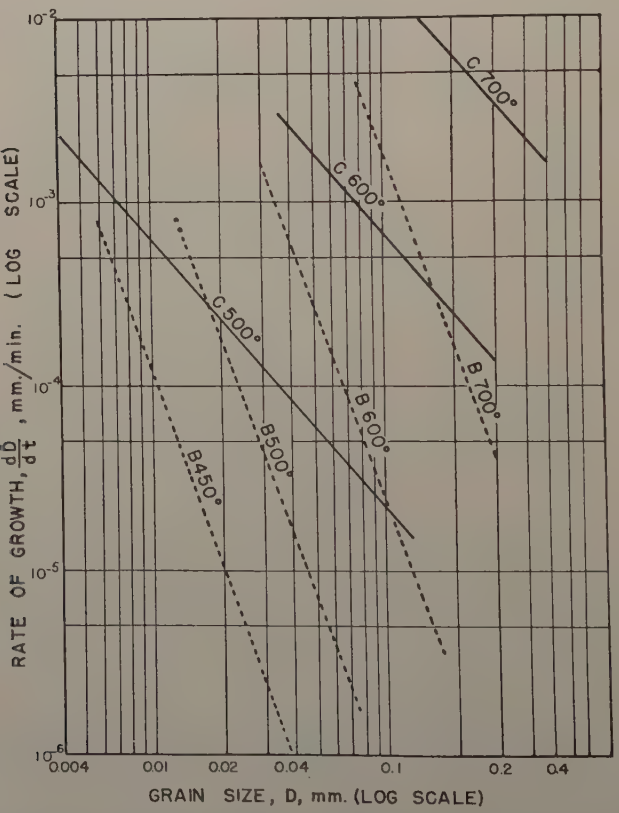
Isothermal grain growth in high purity aluminum. [Ref. (113).]

Fig. 33



Rate of isothermal grain growth in high purity aluminium at 450°c depends only on instantaneous grain size, regardless of whether specimens were previously annealed at 450°c (Nos. 1890 and 1896) or at 400°c (Nos. 1882 and 1886). Data for 400° and 450°c are replotted from fig. 32 for comparison, showing high degree of reproducibility. (Unpublished work by M. L. Holzworth and P. A. Beck.)

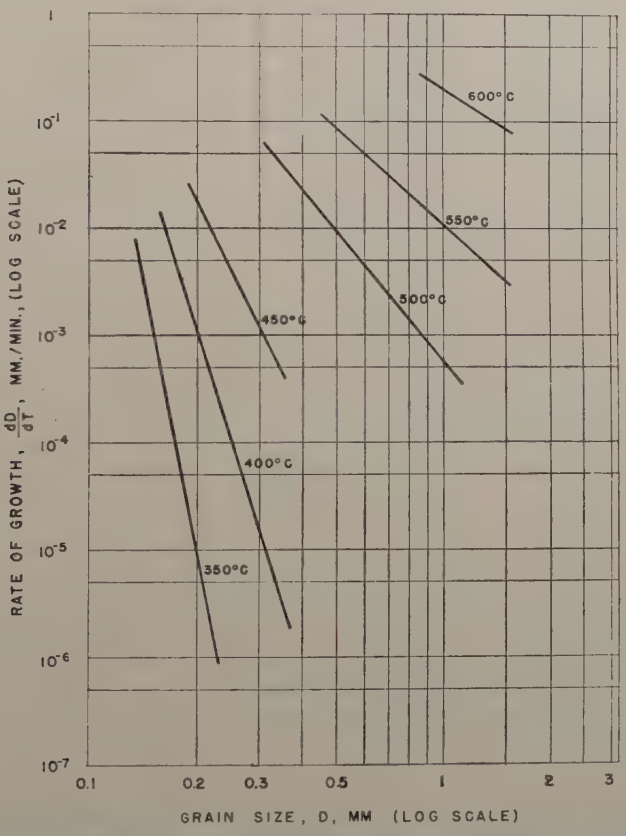
Fig. 34



Average instantaneous rate of grain growth for high purity (————) and for commercial (– – – –) 70–30 brass, as a function of grain size. [Unpublished data by J. Towers and P. A. Beck and from Ref. (114).]

Burke noted[108] that (4) also follows from the assumption that the instantaneous rate of grain growth is proportional to the average grain boundary curvature at the instant. More recently, Machlin[115] showed that the assumption underlying the derivation given above may be justified on the basis of the thermodynamic theory of irreversible processes. Fullman and Fischer[117] obtained experimental evidence for the validity of (5) for bubble growth in soap froth. As stated above, the value of $n=0{\cdot}5$ is also approached in pure brass over a wide temperature range, and in pure aluminium near the melting point. A discussion of the various factors affecting the value of n was reviewed by Fullman[117].

Fig. 35



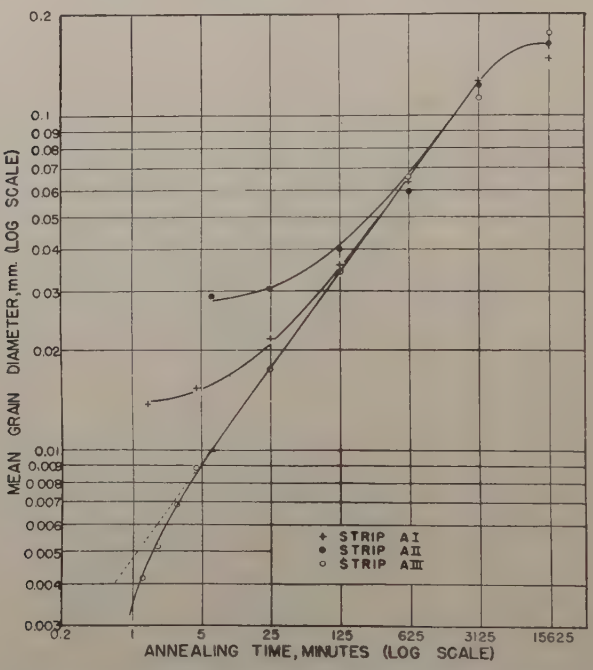
Average instantaneous rate of grain growth for high purity aluminium, as a function of grain size. [Ref. (114).]

It was observed[113] that grain growth on annealing of a rolled polycrystalline metal strip becomes slower than expected, and eventually stops altogether, when the grain size approaches the thickness of the strip (fig. 38). The deviation from the straight line relationship at large grain sizes in fig. 32 results from this effect. This ' specimen thickness

effect ' is also observed in wires, where grain growth stops when the grain boundaries become approximately perpendicular to the wire axis, and the grains occupy the whole cross section of the wire, having an average length of the order of the wire diameter.

The grain growth behaviour discussed above is found only in pure metals or solid solutions.* When finely dispersed particles of a second phase are present in the structure, the grain boundaries tend to cling to these particles[107, 118], fig. 39 (Plate 32), so that grain boundary

Fig. 36

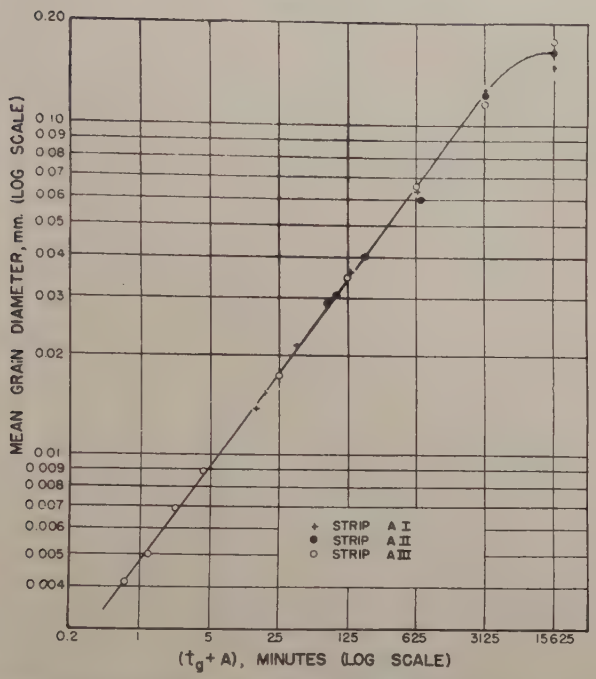


Average grain diameter as a function of total annealing time for high purity 70–30 brass with large penultimate grain size and final rolling of 17·4% (filled circles) with large penultimate grain size and final rolling of 33·3% (crosses) and with small penultimate grain size and final rolling of 85·4% (empty circles). (Unpublished data by J. Towers and P. A. Beck.)

migration is inhibited. In such cases, the rate of grain growth becomes much smaller than without the inhibiting dispersion, as may be seen for instance by comparing the data for commercial brass with those for pure brass (fig. 34). Also the exponent n, i.e. the slope of the logarithmic

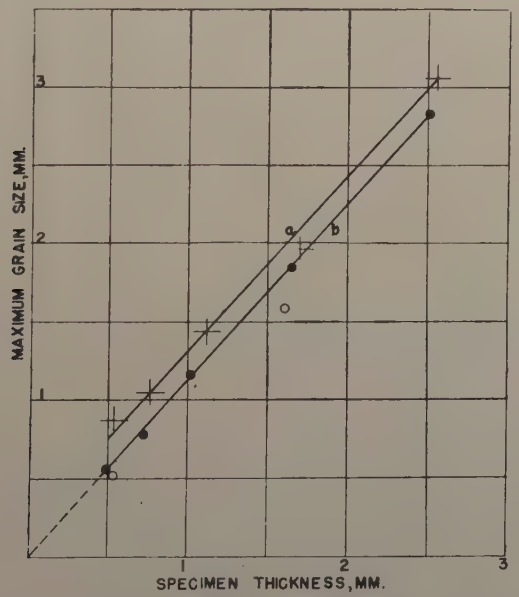
* The solid solutions most extensively studied are alpha brass, as described above, and Al–Mg of various compositions up to 2% Mg[93]. Both of these exhibit gradual grain growth, as that in pure metals. The view has been expressed[116] that solid solution of Ag in Zn inhibits grain growth, but this was based on inadequate experimental evidence, and may be considered unlikely.

Fig. 37



Same grain sizes as in fig. 36, replotted as a function of time for grain growth (i.e. total annealing time *less* time for primary recrystallization *plus* time for imaginary grain growth to produce grain size actually attained by primary recrystallization). (Unpublished data by J. Towers and P. A. Beck.)

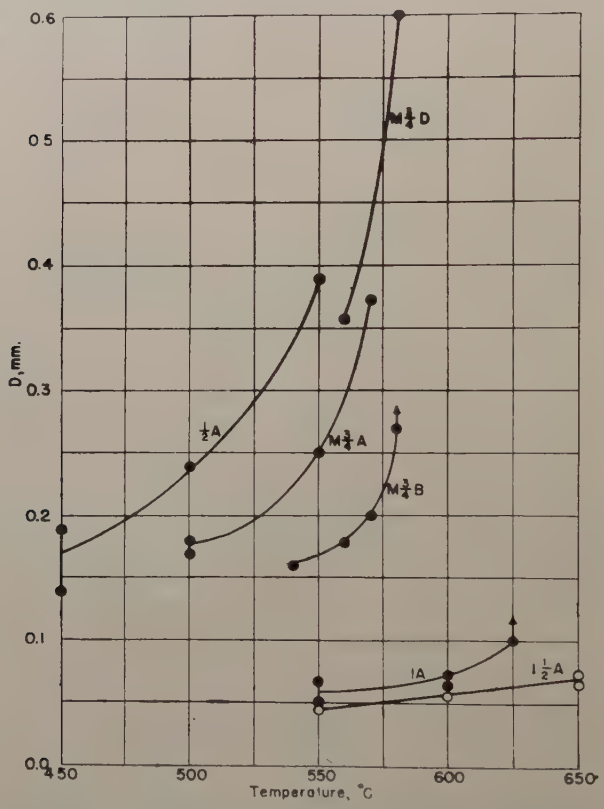
Fig. 38



Maximum grain size attainable by gradual grain growth in high purity aluminium, as a function of specimen thickness, showing 'specimen thickness effect'. (*a*) Specimens not etched before annealing. (*b*) Specimens etched before annealing. Empty circle: specimen extremely deeply etched before annealing. [Ref. (113).]

grain growth lines often decreases; for commercial brass $n=0{\cdot}2$, as compared with $n=0{\cdot}4$ for high purity brass. Still another important effect of inhibition by a dispersed second phase is that grain growth stops completely when a certain 'ultimate grain size' is reached[118]. As the grain size increases during gradual grain growth, the driving energy for further grain growth becomes smaller until, at a certain point, the

Fig. 40

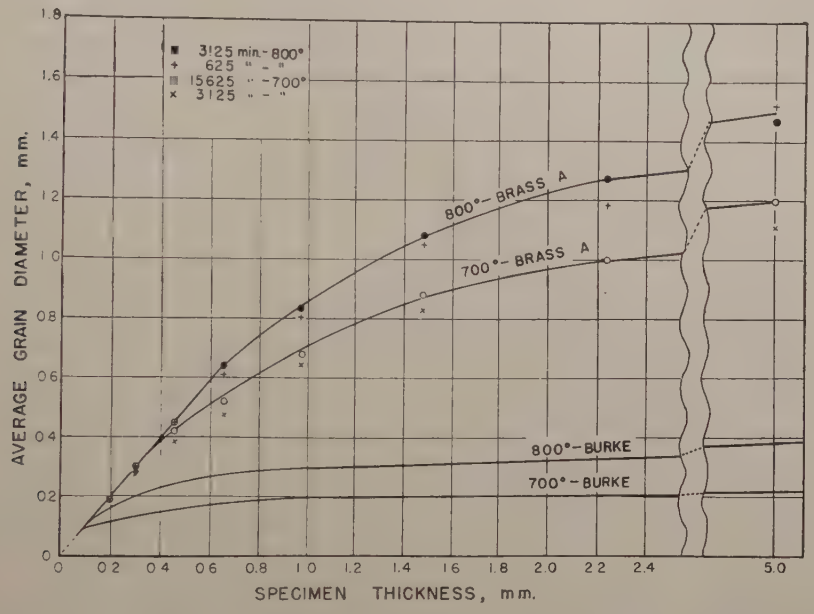


Ultimate grain size as a function of final annealing temperature in Al alloys with 0·6% (½), 0·75% (3/4), 1·12% (1) and 1·65% (1½) Mn, after penultimate anneal at 530°C (A), penultimate anneal at 500°C (B) and penultimate anneal at 600°C (D). For the same final annealing temperature, the ultimate grain size decreases with increasing number of inhibiting particles (i.e. with increasing Mn-content and with decreasing penultimate annealing temperature). [Ref. (118).]

inhibiting effect of the dispersion is just sufficient to prevent further growth[108]. The ultimate grain size so reached is stable over long periods of annealing. As seen in fig. 40, the ultimate grain size decreases with increasing number of dispersed particles. It increases with increasing annealing temperature, although this effect, as observed experimentally[118], may be at least partly due to a decrease in the number of

dispersed particles through their re-solution and coalescence. It is noteworthy that even very few dispersed particles have observable effects on gradual grain growth when the grain size becomes large enough, so that grain growth may be considered as one of the most sensitive indicators of such particles. Inhibition effects have been observed even in high purity brass. Figure 41 shows the somewhat gradual change-over in high purity brass from grain growth limitation by specimen thickness to that by inhibition.

Fig. 41



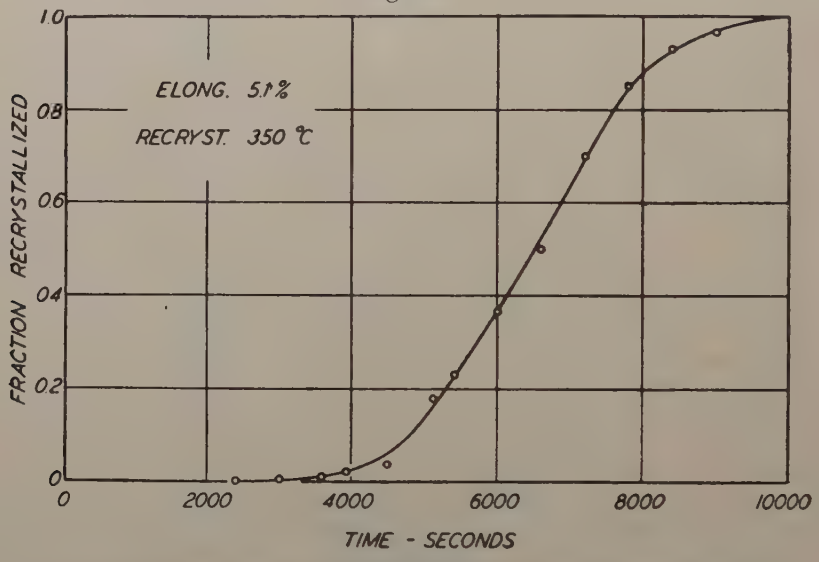
Ultimate grain size in high purity (Brass A) and in commercial (Burke) 70–30 brass, as a function of specimen thickness, for the annealing temperatures indicated. When the thickness is small it limits gradual grain growth, independently of temperature and of purity. For larger thicknesses grain growth is limited by inhibition, so that the ultimate grain size depends on both temperature and impurities. [Unpublished figure, by J. Towers and P. A. Beck, data for commercial brass from Ref. (108).]

(*b*) *Primary Recrystallization*

Studies by Mehl and his associates[119, 120] of the volume percentage of material recrystallized on annealing of slightly extended fine grained polycrystalline specimens as a function of annealing time indicate that the kinetics of the process is typically represented by a 'sigmoidal' curve, as shown in fig. 42. This kinetics is in sharp contrast to that typical for recovery, fig. 6. Johnson and Mehl have shown[121] that the 'sigmoidal' kinetics, where the reaction rate is at first very low, then increases to a maximum and again decreases, is characteristic of transformations that may be formally described in terms of two simultaneous

processes, namely the nucleation of new grains and their growth. Both primary recrystallization and coarsening (secondary recrystallization) may be described in that manner. Goeler and Sachs[122] first carried out a quantitative analysis for primary recrystallization, assuming rates of nucleation and of growth independent of time, and making no allowance for the decrease of transformation rate due to impingement of the recrystallized grains. In the more recent and more complete studies by Stanley and Mehl[119] and by Anderson and Mehl[120] the rate of growth G was still assumed to be constant and the same for all new grains, but no assumption was made regarding the rate of nucleation, N, which was determined experimentally. In these studies, G has been determined by measuring the diameter of the largest grains in a series

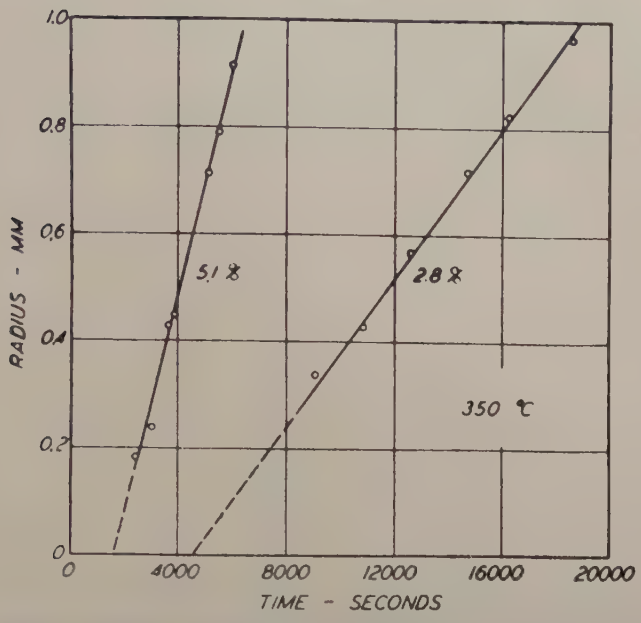
Fig. 42



Isothermal recrystallization curve at 350°c for pure Al extended 5·1%. [Ref. (120).]

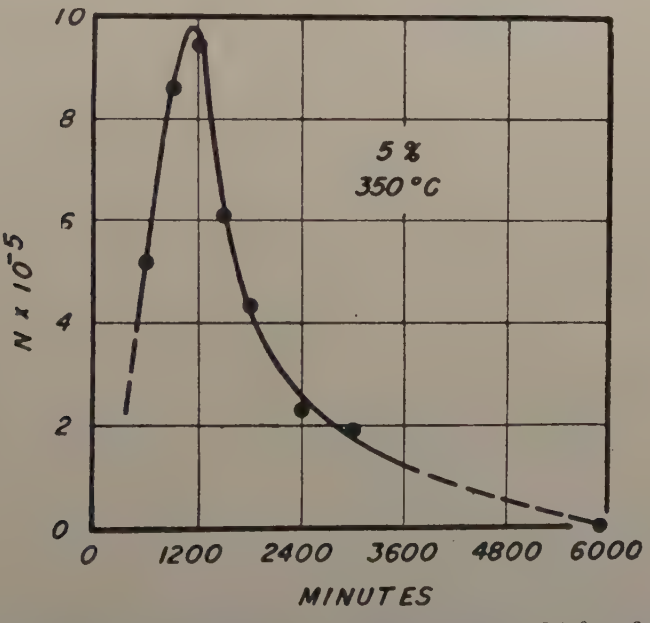
of specimens annealed for various periods at constant temperature. The results fit well the interpretation, according to which the first new grain in each of these specimens formed after the same ' period of incubation ' and then grew at the same rate, so that this rate could be determined from the slope of lines like those in fig. 43, without additional assumptions. After plotting the total number of recrystallized grains as a function of annealing time and graphical differentiation, the rate of increase of the number of recrystallized grains per unit of *untransformed* volume, i.e. the rate of nucleation N, could be calculated as a function of time, since the volume fraction transformed was measured as a function of annealing time. Typical results are shown in fig. 44. Anderson and Mehl determined both the rate of growth and the rate of nucleation in pure Al as a function of deformation and of annealing temperature, figs. 45 and 46.

Fig. 43



Radius of largest grains *vs.* annealing time at 350°c of pure Al specimens, after extensions indicated. [Ref. (120).]

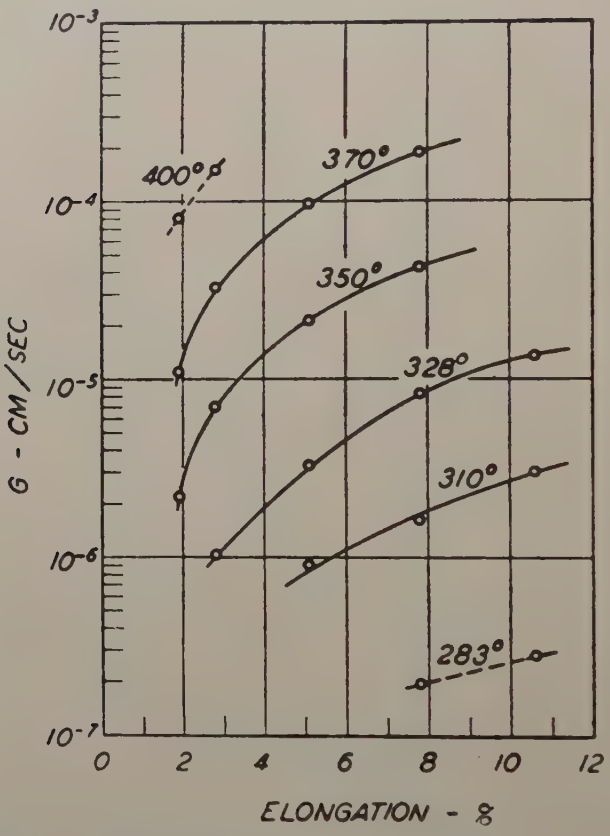
Fig. 44



Rate of nucleation as a function of annealing time at 350°c, after 5% extension. [Ref. (120).]

From the temperature dependence of these rates, activation energies were calculated for both N and G in the range of deformations studied, from about 1% to 15% extension (fig. 47). A proof of the correctness of the interpretation of the process in terms of nucleation and growth and of the assumptions used was obtained by Anderson and Mehl by calculating the amount of material transformed as a function of time

Fig. 45



Rate of growth at temperatures indicated *vs.* per cent extension of pure Al specimens. [Ref. (120).]

from the independently determined values of G and N, with the help of eqn. (6) of Johnson and Mehl[121] for the three dimensional case and of eqn. (7) derived by Stanley and Mehl[119] for the two dimensional case

$$f(t)=1-\exp-\frac{8G^3a}{b^4}\left(e^{bt}-\frac{b^3t^3}{6}-\frac{b^2t^2}{2}-bt-1\right) \quad . \quad . \quad . \quad (6)$$

$$f(t)=1-\exp-\frac{2G^2a}{b^2}\left(\frac{e^{bt}}{b}-\frac{bt^2}{2}-t-\frac{1}{b}\right) \quad . \quad . \quad . \quad . \quad . \quad (7)$$

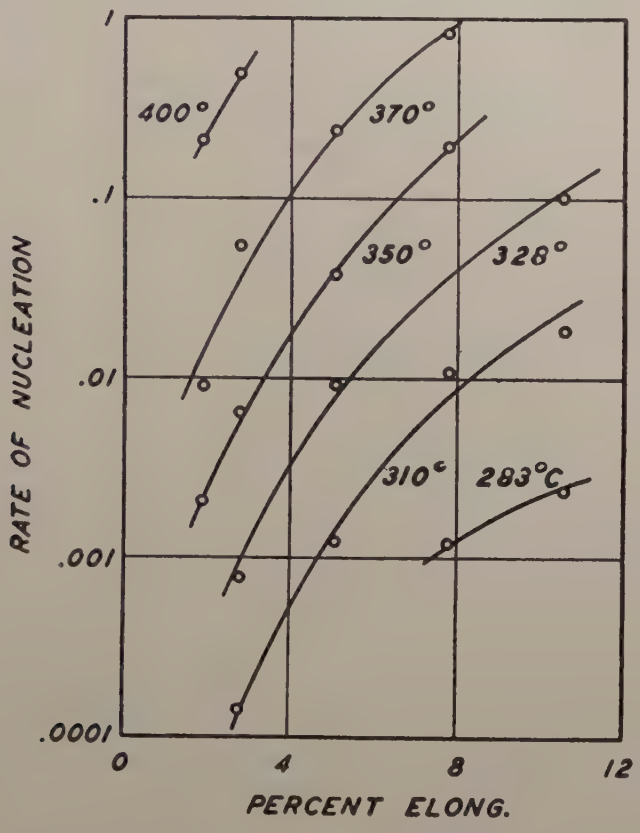
where $f(t)$ is the volume fraction recrystallized at time t and a and b are

the empirical constants giving the rate of nucleation as a function of time according to eqn. (8):

$$N = a \cdot \exp bt. \quad \ldots\ldots\ldots \quad (8)$$

Equation (8) approximates the ascending portion of the curve showing the experimentally determined data in fig. 44. Equations (6) and (7) were derived on the basis of the nucleation and growth picture and of a suitable assumption to take into account the effect on the transformed volume of the impingement of the recrystallized grains on each other.

Fig. 46



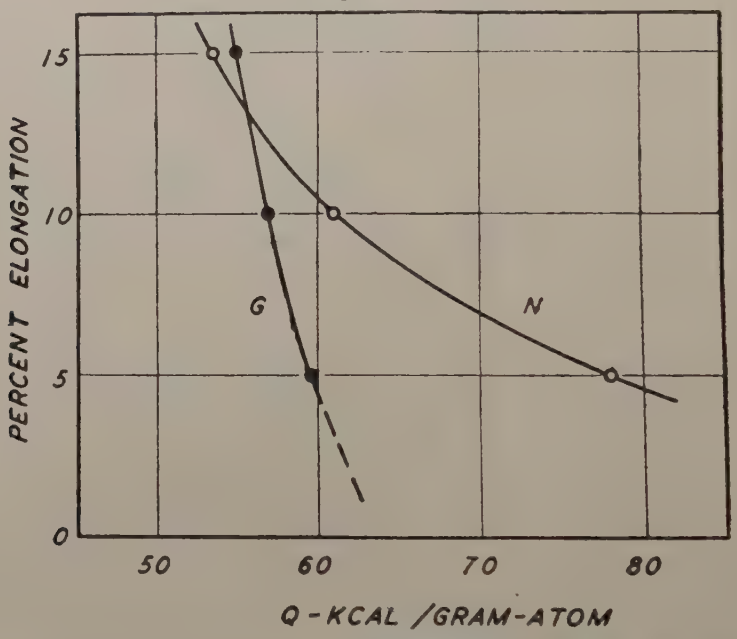
Rate of nucleation at temperatures indicated at 20% recrystallization *vs.* per cent extension of pure Al specimens. [Ref. (120).]

The curves so calculated were in good agreement with the values of the volume fraction transformed, directly measured as a function of annealing time.

In a recent paper, H. Suzuki[123] indicated that in high purity aluminium the rate of growth G of recrystallized grains after small extensions decreases with gradual grain growth in the deformed matrix material. This observation was interpreted by Suzuki in accordance with the view expressed

earlier[11] that the grain boundary interface energy may contribute a substantial part of the total driving energy for primary recrystallization in a polycrystalline metal, when the matrix grain size and the deformation are very small. In addition, the residual stored energy of cold work itself may be gradually eliminated by continuous grain growth[124, 109]: the portions of the matrix swept by moving boundaries become essentially free of strain and of substructure. For both of these reasons, the driving energy for the growth of large recrystallized grains may be expected to decrease as a result of the concurrent gradual grain growth in the matrix. Suzuki's results for low deformations, as shown in fig. 48, differ radically from those of Anderson and Mehl, fig. 43. The cause of

Fig. 47

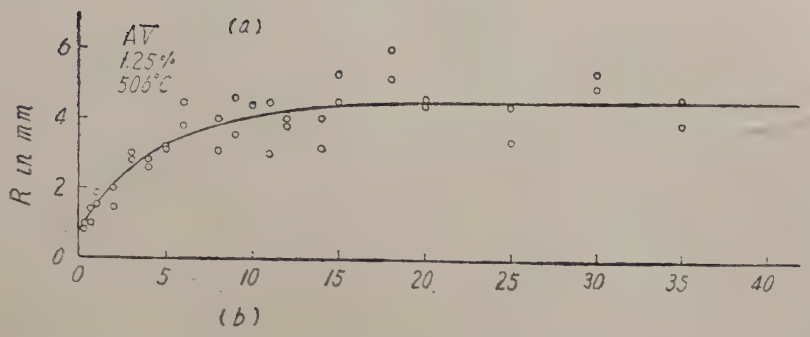


Activation energies for nucleation and growth in pure Al *vs.* per cent extension. [Ref. (120).]

the discrepancy may be the higher purity of the aluminium used by Suzuki. The possibility that the metal employed in the work of Anderson and Mehl may have contained dispersed second phase particles in sufficient quantity to inhibit continuous grain growth in the matrix is suggested also by the small matrix grain sizes reported by these investigators.

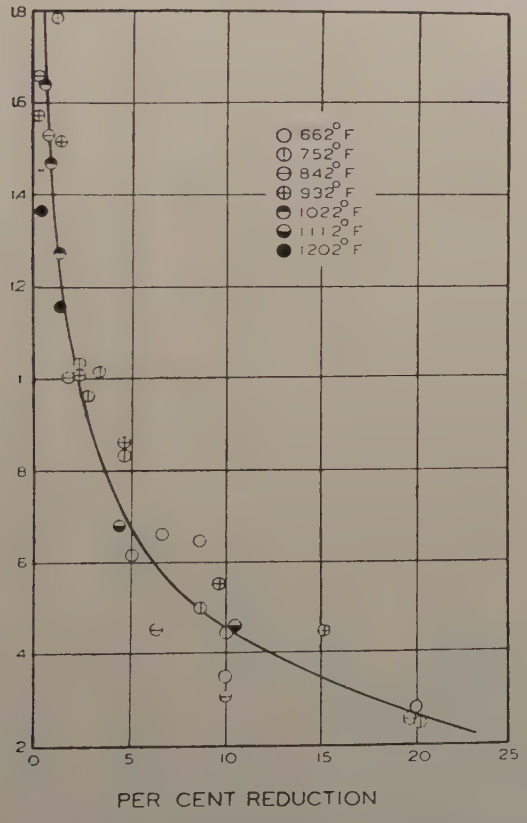
With increasing deformation, the number of recrystallized grains becomes larger, that is the grain size smaller[99] (fig. 49), until the new grains can no longer be separated from the original deformed grains by means of simple criteria, such as size. Also, a considerable fraction of the total growth of each new grain may take place within a single deformed matrix grain, and the rates of nucleation and growth vary greatly with

Fig. 48



Radius of grains artificially nucleated *vs.* annealing time in minutes at 506°c, after 1·25% extension. [Ref. (123).]

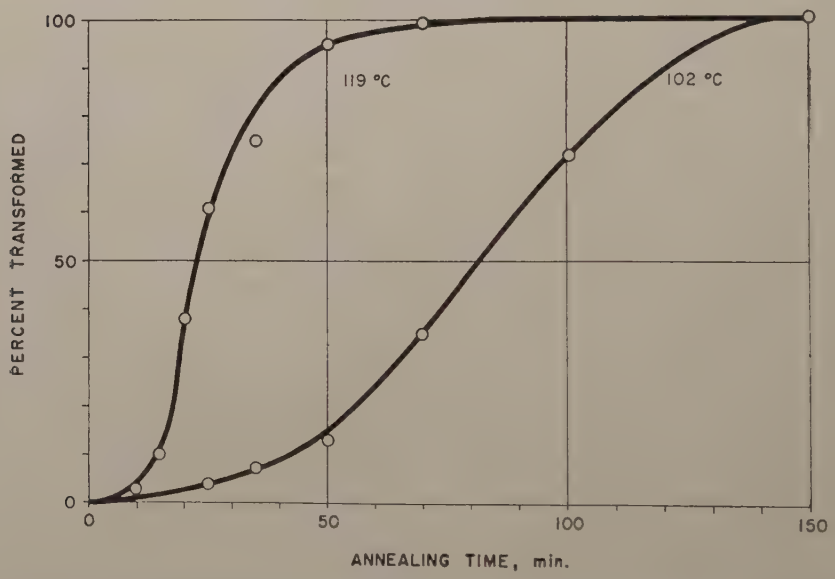
Fig. 49



Average grain size (mm) as recrystallized at temperatures indicated, after varying plastic deformation by extension. (Eastwood, James and Bell.)

the orientation of the individual matrix grains with respect to the stress axes. Because of this variation and of that resulting from the inhomogeneity of the plastic deformation due to grain boundaries and kink bands, the local differences in the rates of nucleation and growth are very great. For the various reasons mentioned a detailed analysis of nucleation and growth rates, such as that accomplished by Mehl and his associates for primary recrystallization after small deformations (where the described local variations are averaged out due to the large size of the recrystallized grains), would be difficult or impossible to carry out for metals more severely deformed. However, the average overall kinetics, that is the volume fraction of material recrystallized as a function of annealing time, can be determined accurately even after severe deformation in cases where recrystallization is associated with a pronounced

Fig. 50



Isothermal recrystallization curves for pure Cu at temperatures indicated after 98% cold rolling. [Replotted from Ref. (126).]

reorientation of the texture. Decker and Harker[126] were thus able to measure quantitatively the kinetics of reorientation in 98% rolled copper, where the recrystallized material has a very sharp 'cube texture'. Figure 50 shows that the kinetics is typically 'sigmoidal' like that found after small deformations (fig. 42). Similar results were obtained by Seymour and Harker[127] for a highly rolled Ni–Fe alloy, and by McGeary and Lustman[89] for zirconium. These results suggest that primary recrystallization even after severe cold working may be described as a nucleation and growth process, a conclusion which is, of course, in excellent agreement with numerous qualitative microscopic observations.

(*c*) *Coarsening* (*Secondary Recrystallization*)

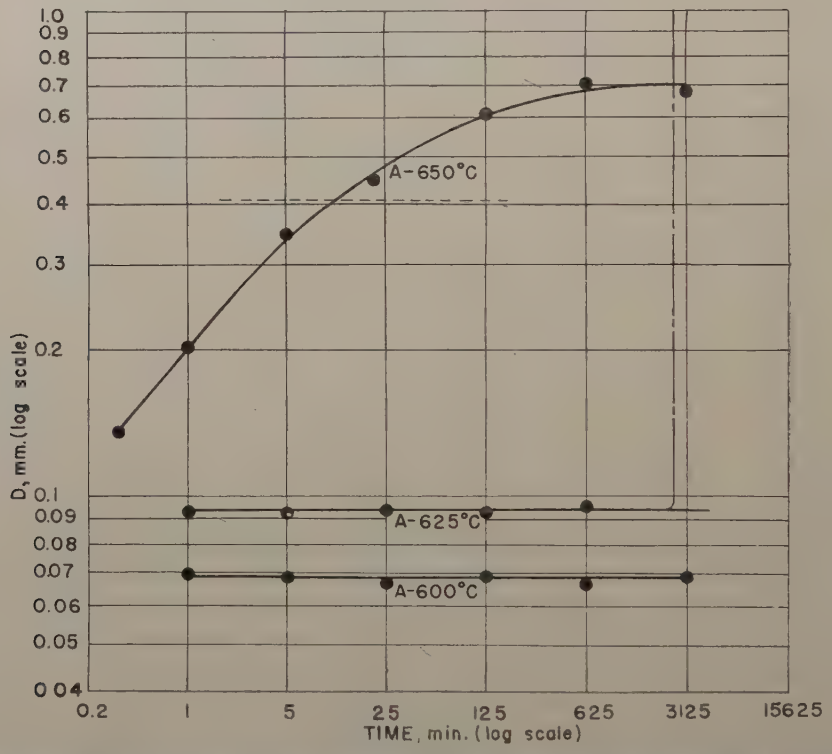
As mentioned above, grain growth in soft annealed metals may take place by a gradual or 'continuous' mechanism, which leaves the type of grain size frequency distribution relatively unchanged. However, if conditions are such that most grain boundaries in the structure are not able to migrate, although there are a few which are free to do so, the kinetics of the process of grain growth become quite different from those discussed in (*a*). The relatively few growing grains become very much larger than the surrounding matrix grains (fig. 51, Plate 33), so that the type of grain size frequency distribution is also radically altered. The splitting of the structure into two clearly separated portions, namely the new, coarse grains and the remaining relatively unaltered matrix grains, gave rise to the designation 'duplex structure'. There are now known[129] two fundamentally different causes that may be responsible for the immobility of grain boundaries and, hereby, for the type of grain growth discussed.* One of these is inhibition by dispersed particles of a second phase, as described above. The other one is connected with the orientation dependence of boundary mobility, to be discussed in § 4.4; if the orientation of almost all grains in the structure is nearly the same, most grain boundaries have very low mobility, and only the few grains with deviating orientations can grow. These two causes of boundary immobility may be operative independently from each other, but they are in many technological situations both active at the same time. For this reason, it is not always easy to tell whether a given case of coarsening is of the inhibition-dependent or of the texture-dependent type. Also in some cases the dispersed particles are so small that they can not be detected microscopically. For instance, the well known coarsening phenomena in Al-killed low carbon steels are typical of inhibition-dependent coarsening, yet for many years it has not been possible to identify the dispersed particles responsible for them. Recent investigations indicate that this dispersed phase is connected with the nitrogen content, so that it is most likely aluminium nitride[128].

A detailed investigation[118] of grain growth in Al–Mn alloys containing a dispersed second phase showed that coarsening occurs when a small initial grain size is stabilized by precipitating a large amount of inhibiting second phase and the structure is then annealed for prolonged periods of time at a temperature just below the solvus point (fig. 52). The latter treatment causes the precipitate to redissolve partially and to coalesce, so that the number of inhibiting particles is slowly reduced to a lower value. In the course of annealing, the inhibition initially sufficient to immobilize the grain boundaries at a small grain size level is gradually decreased, until at certain localized areas, where conditions happen to be particularly favourable, grain growth may start, while it is still inhibited

* The third known cause of grain growth stoppage, namely the 'specimen thickness effect', is not known to give rise to coarsening.

elsewhere. Once a grain becomes considerably larger than its neighbours it can grow at their expense at a level of inhibition still sufficient to prevent the migration of boundaries between the small grains. This condition obtains, since at the edges of the large grains the surface tension forces of the intersecting boundaries are further removed from equilibrium, and also the boundaries of the larger grains have greater curvatures than those between the small matrix grains[130]. It is significant that, if the inhibition is removed too fast at too many points, for instance by annealing at a temperature slightly above the solvus point, the grain

Fig. 52



Isothermal grain growth at final annealing temperatures indicates, in Al+1·1% Mn alloy *vs.* annealing time. Vertical dash-dot line indicates approximate start of coarsening. [Ref. (118).]

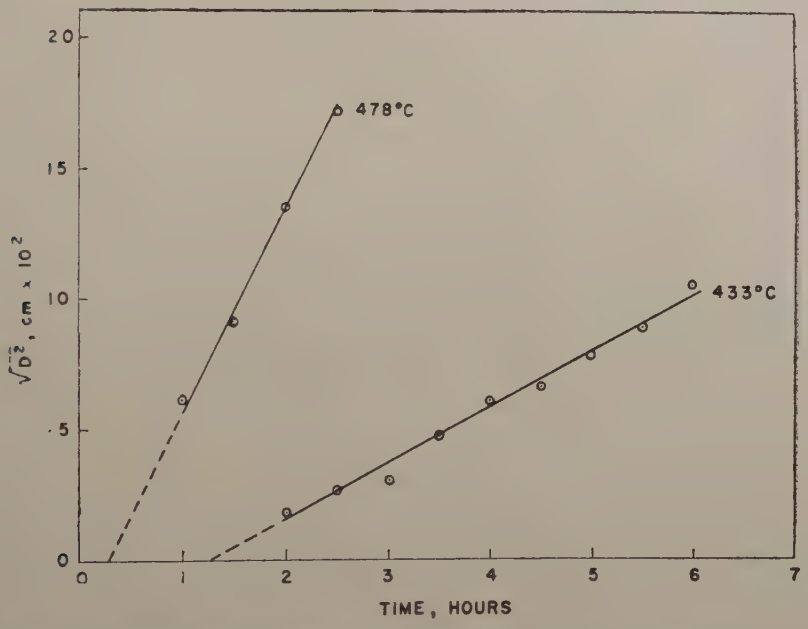
growth then taking place is of the gradual rather than of the discontinuous type. Since gradual grain growth is limited by the specimen thickness, while the grain size attained in coarsening is not, it is possible under suitable conditions to arrive at much larger grain sizes by annealing at a somewhat lower temperature than those developed at a higher temperature (see fig. 52).

Texture-dependent coarsening occurs only when the matrix has a well developed single orientation texture. If the matrix texture has more than one component, no matter how sharply each is developed, in general

the structure contains many high angle boundaries and these are free to migrate, so that the resulting grain growth is of the gradual type. The only examples of pure texture-dependent coarsening now known are those occurring in a cube texture matrix, or in a matrix obtained by primary recrystallization from a homogeneously deformed single crystal.

The kinetics of coarsening has not yet been investigated in as much detail as that of gradual growth and of primary recrystallization. The early investigations of W. Feitknecht[131], Karnop and Sachs[132] and others established the general features of the phenomenon, and showed that its kinetics may be similar in many respects to a nucleation and growth type process. This was later demonstrated particularly clearly

Fig. 53

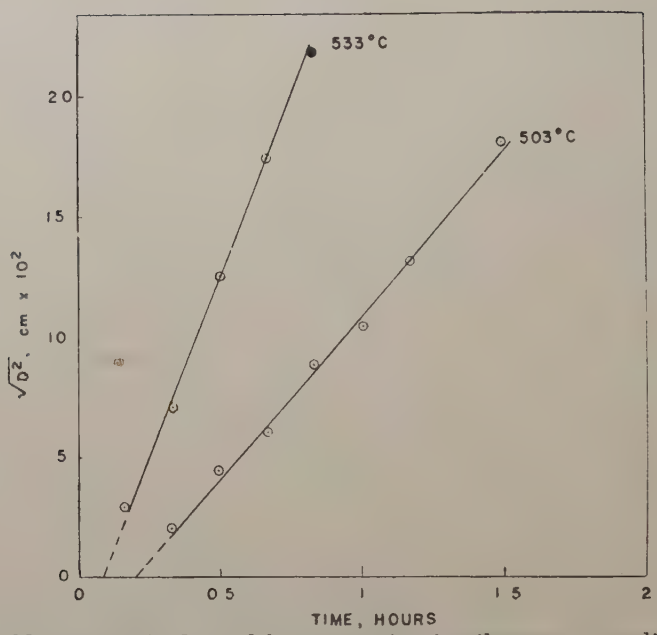


Diameter of largest grain formed by coarsening in silver *vs.* annealing time at temperatures indicated, after final rolling reduction of 50%. [Ref. (133).]

by the investigations of Rosi, Alexander and Dube[133]. The analogy between the kinetics of primary and secondary recrystallization was emphasized by Burke and Turnbull[10]. Although there are so far no quantitative data available on the volume transformed as a function of isothermal annealing time, in analogy to the relationships in primary recrystallization a ' sigmoidal ' type transformation curve is certainly to be expected, at least when the specimen size is sufficiently large in comparison with the resulting grain size.

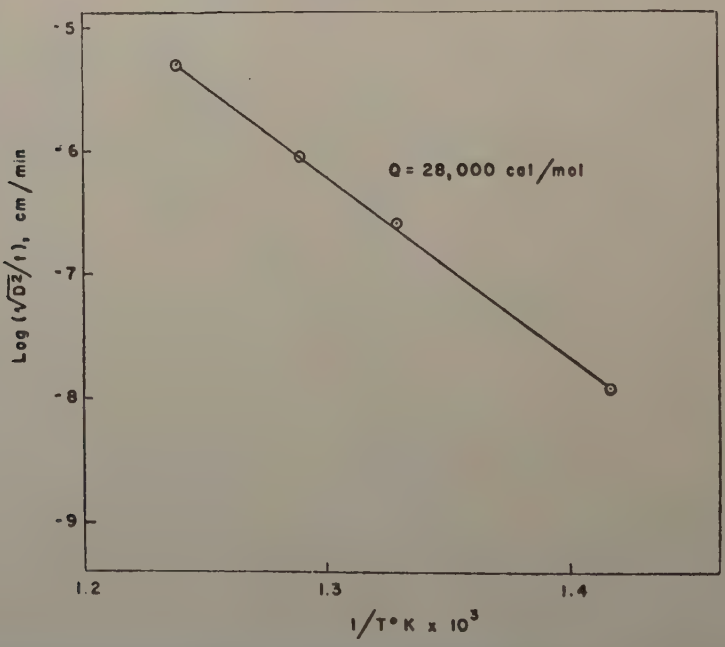
The best published data on the rate of growth in coarsening are those of Rosi, Alexander and Dube for silver, shown in figs. 53, 54, and 55. As seen in fig. 55 these data define a thermal activation type temperature

Fig. 54



Diameter of largest grain formed by coarsening in silver *vs.* annealing time at temperatures indicated, after final rolling reduction of 50%. [Ref. (133).]

Fig. 55



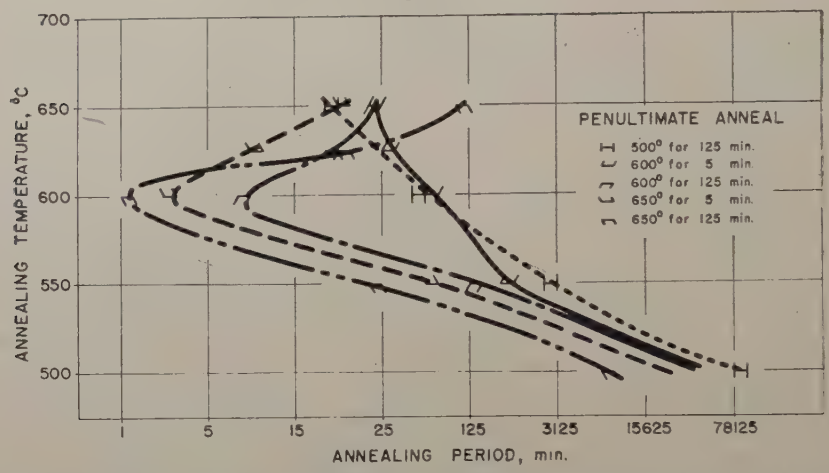
Logarithm of growth rate *vs.* reciprocal absolute temperature for coarsening in silver. [Ref. (133).]

dependence, with an activation energy of 28 kcal/mole. According to the calculations of Turnbull[134] the free energy of activation for this boundary migration is 21·5 kcal/mole as compared with 18·6 kcal/mole for the free energy of activation for self diffusion along high energy grain boundaries. The reasonably good agreement here supports the interpretation[55] according to which the elementary process determining boundary mobility is similar to that determining the rate of self diffusion along the boundary. Certain theoretical questions involved are discussed fully in the review by Burke and Turnbull[10], which also gives a tabulation of available data for free energies and entropies of activation, calculated on the basis of the absolute rate theory.

Although statistical rates of ' nucleation ' have not been determined for coarsening, Rosi, Alexander and Dube measured incubation periods for individual secondary grains in silver as a function of temperature. They found a temperature dependence of the reciprocal incubation period very similar to that of the rate of growth in the same material, with an activation energy of 30 kcal mole. From this, they concluded that the incubation period corresponds to ' invisible growth ' at a lower rate when the secondary grain is not sufficiently larger than the neighbouring matrix grains and it has not yet emerged as a secondary grain. They found that in order to be able to grow even very slowly, a secondary grain had to be at least twice as large as its neighbours, and that its rate of growth increased with its size up to a grain diameter approximately 13 times larger than the average matrix grain size. Beyond that size, the rate of growth became independent of the size of the secondary grain. Even though the material used in this investigation was high purity silver, the occurrence of distinct coarsening in a matrix containing several orientations suggests that second phase inhibition may have played a role. (In 70–30 brass, under analogous conditions of preferred orientation, gradual grain growth instead of coarsening takes place.) For inhibition-dependent coarsening, the measured incubation periods are often influenced by such factors as the rate of re-solution of the dispersed phase, and the extent of gradual grain growth in the matrix prior to coarsening. A rather clear indication that even small differences in matrix grain size may cause large changes in the incubation period in coarsening was obtained in the investigation with Al–Mn alloys[118]. Complex relationships sometimes encountered are shown in fig. 56, from a study by Sperry and Beck[136] of coarsening in a high purity Al-alloy with 0·35% Fe. When a high penultimate annealing temperature was used, the incubation period showed a clear-cut minimum at a final annealing temperature of 600°C ; at higher final annealing temperatures the incubation period actually increased, so that the apparent ' activation energy ' between 600°C and 650°C was *negative*. This was not the case when the penultimate annealing was carried out below 600°C, or at 600°C for 5 minutes only. Apparently, the amount of Fe dissolved during the penultimate anneal and re-precipitated during cold working

had an important part in the overall inhibition effect. Presumably, the fast re-solution of this very fine precipitate on annealing in the temperature range of 600°–650°c allowed some gradual grain growth to take place in the matrix and thus delayed the ' nucleation ' of coarse grains.

Fig. 56



Approximate annealing periods for beginning of coarsening in pure Al—0·34% Fe alloy at final annealing temperatures from 500° to 650°c, after the various penultimate annealing conditions stated and a final rolling reduction of 80%. (Unpublished research by P. R. Sperry and P. A. Beck.)

4.4. *Problems of Nucleation in Recrystallization*

As mentioned in the previous section, it is commonly recognized that primary recrystallization, as well as secondary recrystallization (coarsening) may be at least formally described in terms of nucleation and growth, although this is not true in the case of gradual grain growth. Even though the growth of new grains after a period of incubation is readily observed by metallographic means in both primary recrystallization and in coarsening, the origin of the ' nuclei ' of these new grains has been the subject of much speculation and it is still not definitely known.

(*a*) *Primary Recrystallization*

All theories recognize the condition, first formulated by R. Becker[137], that a strain-free nucleus is stable and capable of growth in a strained matrix if it is not smaller than a certain critical size. At the critical size, for a small amount of growth the free energy gained from the additional volume transformed is just sufficient to compensate for the free energy required to increase the interface area. The various theories differ mainly in regard to the mechanism by which nuclei of critical or larger size are formed and to the interpretation of the incubation period. In their recent review, Burke and Turnbull[10] gave a rather complete description of the various theories that have been proposed to explain nucleation in primary recrystallization.

In the ' conventional nucleation theory ' strain-free nuclei are formed by means of thermal activation, at a rate increasing sharply with the local lattice strain. The incubation period in this theory is the time required at a given temperature for a sequence of thermal fluctuations until a sufficiently powerful activation event is encountered leading to the formation of a strain-free ' nucleus ' of not smaller than critical size at a certain site in the deformed matrix. Burke and Turnbull estimated that a local elastic strain of the order of 0·2 or greater would be required in order that a certain site should effectively contribute to the formation of recrystallization nuclei by this mechanism. It is necessary to assume that such large strains are present in at least a small fraction (for example in the order of 10^{-6}) of the total volume. If this volume fraction is small enough, the locally very high energy of these sites and the relatively high energy regions which it is necessary to assume in their immediate vicinity in order that the nuclei can grow, may constitute only a very small part of the residual stored energy of cold work in the specimen as a whole. Consequently, it is possible to postulate a strain distribution in the cold worked matrix that would allow nucleation by this mechanism, without requiring more average stored energy than the measured values. However, it is uncertain what fraction of the measured stored energy of cold work is actually retained at the point in the annealing process where recrystallization starts. As stated in § 4.2, there are at least some instances where apparently most of the stored energy of cold work is released prior to the main part of recrystallization. It is, therefore, probable that this theory is not generally applicable.

In the so-called ' block theories ' the nucleus is pictured as a pre-existing block of a size sufficient to grow, that has either remained during the deformation process relatively strain-free with respect to its surroundings (' low energy block theory '), or has become relatively strain-free during annealing by means of some process that decreases the strain uniformly in the whole block (' high energy block theory '). In contrast to the ' conventional nucleation theory ', in the ' high energy block theory ' the nucleus-block may be quite large, since the underlying assumption is that the nucleus is not formed by a single activation event, but rather by a gradual process requiring a large number of thermal activation events within the block volume, whose effect is cumulative.

In the low energy block theory, the incubation period is the time for initial slow growth before the higher, final growth rate is attained. In analogy to the findings of Rosi, Alexander and Dube[133] for coarsening, it has been recognized that in primary recrystallization, too, the rate of growth of a nucleus of the critical size must be practically zero, and that this rate may be expected to increase gradually with increasing size, as the gain of free energy for a small transformed volume becomes considerably larger than the loss due to the corresponding increase of interface area. Recently, Cottrell[138] proposed an additional reason for low initial rates of growth, and for increased rates as the new grains become

larger. As discussed in the following section, there is now good evidence indicating that the mobility of a grain boundary depends on the orientation difference between the two grains on either side of it, having very low values for small disorientations. Cottrell suggests that originally the nucleus may be only slightly disoriented with regard to its environment, but that as it grows, the boundaries reach nearby matrix regions (e.g. kink bands) where the orientation is quite different. From there on the rate of growth may become much larger. In addition to giving a plausible interpretation of the incubation period, this picture also has the advantage of permitting a considerably smaller size for a stable nucleus, in terms of the Becker condition, than the assumption of a nucleus of strongly deviating orientation, the radius of a stable spherical nucleus is proportional to the interface energy[137], which in turn increases greatly with the disorientation[139].

The theory of nucleation proposed by Cahn[140], assumes that essentially strain-free blocks of very small size form by polygonization in the early phase of annealing. It is implied, however, that the incubation period is not determined by this process, but rather by the rate at which some of the resulting subgrains grow sufficiently large at the expense of neighbouring subgrains of nearly the same orientation to serve as nuclei capable of growing at the expense of matrix material of strongly deviating orientation. In this theory, which in part leans on the earlier suggestion by Burgers[141], the incubation period corresponds to the annealing time necessary for the initial subgrain growth.

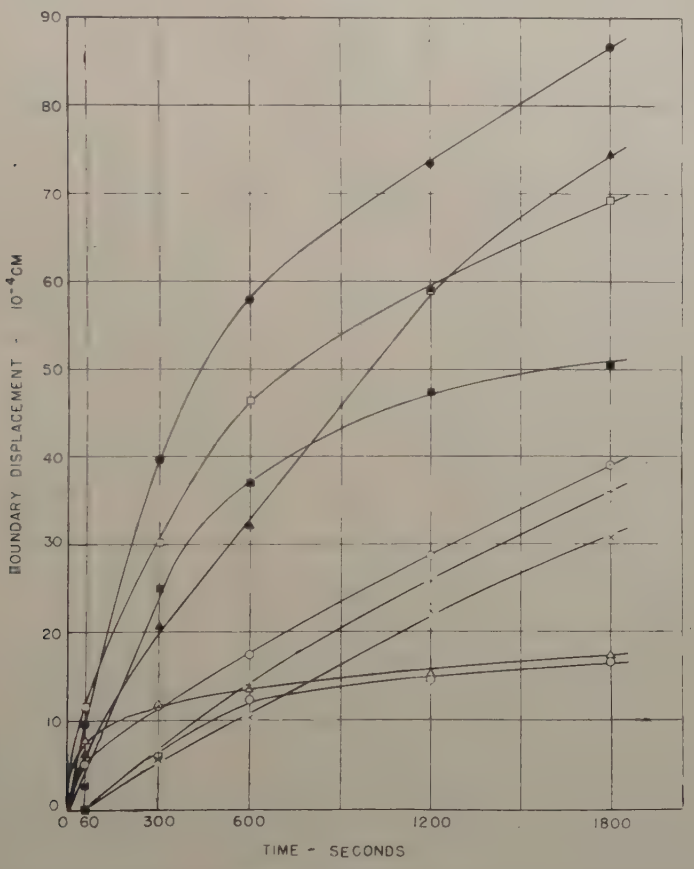
Since thus polygonization itself is not an essential part of the theory, the usefulness of Cahn's idea is not affected by the recent experimental confirmation of the picture originally proposed by Bragg[142] and supported by Burgers[143], according to which the subgrains are essentially formed during the deformation process itself, rather than during subsequent annealing.* In view of microscopic observations of subgrain behaviour[145] and of the occurrence of recrystallization in practically strain-free material, it appears that Cahn's theory, together with the mechanism proposed by Cottrell, represents the best basis available at present to describe the formation of recrystallization nuclei in deformed single crystals and in severely deformed polycrystalline material.

Microscopic observations by Sperry and Beck[146] showed that in polycrystalline high purity aluminium, after deformation of less than approximately 40% by rolling, the formation of new nuclei on annealing is relatively rare and that volume elements relatively free of substructure form instead by the 'strain-induced migration' of already existing boundaries. Figure 57 (Plate 33) shows such an area, that became largely free of substructure when the boundary migrated from its original position 1 to a new position 2. The microscopic technique used (oxide film and sensitive tint illumination) indicated that the area between 1

* On the other hand, the theory of nucleation proposed in 1949 by Beck[186] was based on polygonization, and it should be, therefore, abandoned.

and 2 has the same orientation as that of the original grain A, from which the boundary moves away. This conclusion was confirmed by means of back reflection Laue patterns. Measurements of the rate of this kind of boundary migration were carried out[11] by an annealing treatment interrupted at suitable intervals and by developing surface markings showing the successive boundary positions (fig. 58, Plate 34). Within the accuracy of the measurements, no incubation period could be detected ; growth started immediately upon the immersion of the specimens in the

Fig. 59



Isothermal strain-induced migration at 350°c of several boundaries in the pure Al specimen shown in fig. 58. [Ref. (11).]

molten salt bath at the annealing temperature, and in most cases the initial rate of boundary migration was the highest (fig. 59). The volumes essentially free of substructure so formed eventually attain dimensions commensurate and even larger than the matrix grain size, and on further annealing, they begin to absorb other surrounding grains.

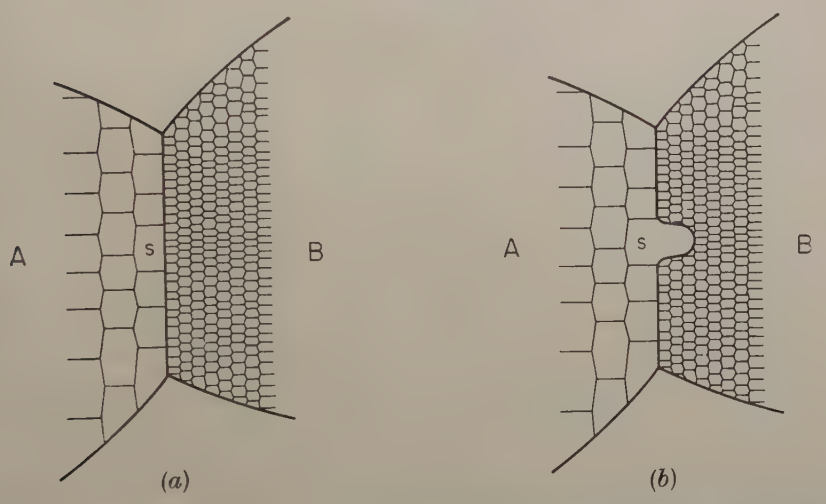
It seems reasonable to assume that under conditions similar to those

used in Anderson and Mehl's[120] experiments, that is when both the initial grain size and the strain is small, this process represents the predominant ' nucleation ' mechanism. This conclusion is confirmed by the findings of Tiedema[148], who has shown by means of x-ray diffraction techniques that the central area (' nuclear spot ') of large grains produced in pure Al strips by the strain-anneal method often consists of many subgrains of an orientation similar to the main portion of the grain. This is just what may be expected if the large grain is nucleated by the mechanism described, since the recrystallized grain can not absorb the deformed grain (A in fig. 57) of nearly the same orientation from which it originates[135]. Tiedema also observed that the main portion of the large recrystallized grains often consists of two or more slightly disoriented domains. As noted by Crussard[12], this, too, is in agreement with the proposed mechanism, since it has been observed microscopically[146] that the relatively substructure-free area swept by the migrating boundary (fig. 57) in effect usually does contain a few low angle boundaries. As seen in fig. 58, the origin of these boundaries may be traced back to independent nucleation along the original grain boundary of several ' buds ' of nearly the same orientation, which later meet and become joined by low angle boundaries. Apparently, each ' bud ' originates in a single subgrain, S, of deformed grain A (fig. 60 (*a*)), directly adjacent to the boundary of deformed grains A and B. If this boundary is mobile by virtue of the large disorientation between A and B and, if opposite to subgrain S of grain A the subgrain size in grain B is much smaller, so that there is an unbalanced ' pull ' in the direction of B, the conditions are favourable for the process to start immediately on annealing. In the initial stage of the boundary migration, shown in fig. 60 (*b*), the driving energy is the residual portion of the stored energy of cold work (which may be merely the energy of the substructure) in grain B less the energy required to extend the migrating high angle grain boundary. After the linear dimensions of the substructure-free volume become approximately twice as large as the average matrix grain size, the driving energy is increased by the ' pull ' of the high angle grain boundaries between adjacent matrix grains, which in itself now compensates for the energy required to expand the migrating boundary. Consequently, one may assume a period of relatively slow growth during the initial stage, when the net driving energy is low, and subsequent fast growth after the ' recrystallized ' grain becomes several times larger than the matrix grains. On this picture, the ' incubation period ' would correspond to the initial stage of slow growth. The peculiar variation of the rate of nucleation with annealing time, as reported by Anderson and Mehl[120], fig. 44, could then be interpreted as the frequency distribution for the periods of slow growth for the various ' nuclei '.

Both the mechanism of ' nucleation ' just described, and the theory of Cahn discussed previously, have in common that they do not presuppose any lattice stress other than that implied by the presence of subboundaries. In both cases, the site of ' nucleation ' is neither a high stress point

nor a low stress point, but rather it is characterized by relative freedom from substructure. According to the Cahn–Cottrell picture, the substructure-free volume of sufficient size develops on annealing by subgrain growth, and the high energy boundary with the matrix develops gradually as the disorientation increases during the slow initial growth. In the other mechanism, both the subgrain of sufficiently large size with respect to its neighbours and the high angle boundary are present already in the initial state, and the migration of the high angle boundary may start immediately. The apparent incubation period observed here macroscopically, for instance by Anderson and Mehl[120], has a rather different physical significance from the incubation period in accordance with

Fig. 60



Schematic drawing of grains A and B with their substructure (*a*) before, and (*b*) after the beginning of strain-induced boundary migration. This boundary migration is ' nucleated ' by subgrain S of grain A, which is particularly large and is directly adjacent to the grain boundary.

the Cahn–Cottrell theory. The ' strain-induced boundary migration ' mechanism is apparently prevalent at low to medium deformations, and it certainly requires a polycrystalline matrix. Presumably, its prevalence may increase with decreasing matrix grain size, since with a larger number of grain boundaries favourable combinations of subgrain size contrast and disorientation should become more frequent. On the other hand, the Cahn–Cottrell mechanism may predominate in the case of single crystals, or of polycrystalline metal severely deformed, where most of the recrystallized grains originate within the deformed grains, rather than directly at the boundaries. The frequency of coincidence of the required local gradients of subgrain size and of orientation may be expected to increase with increasing deformation.

Microscopic observations[145] with the oxide film-polarized light technique revealed the fact that, at least in severely rolled high purity

Al of large grain size, subgrain growth may lead directly to structures that under the microscope appear very much like recrystallized material. Under such conditions, the subgrains were observed to grow to sizes of the order of 0·02 to 0·03 mm ; they can not then be distinguished from ordinary grains by their size. Since the disorientation among neighbours tends to increase with progressive subgrain growth, this criterion also becomes uncertain. It is quite possible that in some cases subgrain growth may in this manner change over imperceptibly into what might be designated as gradual grain growth, leading to even larger grain sizes, without ever passing through a discernible stage of nucleation and growth, characteristic of primary recrystallization.

(*b*) *Coarsening* (*Secondary Recrystallization*)

Since, as it is now quite generally recognized, coarsening does not depend on the presence of lattice imperfections in the matrix[147], the Becker condition must be interpreted entirely in terms of interface energies. Accordingly, a stable, growing nucleus must be large enough so that the surface energy gained by shortening the interfaces between the matrix grains adjoining the nucleus must be at least equal to the surface energy lost by the corresponding extension of the nucleus-matrix interface.

In the case of coarsening in rolled and recrystallized silver, Rosi, Alexander and Dube[133] found that the nucleus must be at least twice as large as the average matrix grain size. When the matrix has a single orientation texture, so that most of the matrix grain boundaries have low energies, while the interface between the coarse grain and the matrix consists of high energy boundaries, as in the case of coarsening in cube-texture copper, the nucleus must be presumably much larger. Such considerations practically exclude all possibilities of nucleation by thermal activation, even if it is assumed that nuclei form at the intersections of twin-boundaries, grain boundaries and specimen surface, as proposed by Kronberg and Wilson[149].* The only reasonable picture then is that the 'nucleus' for coarsening is a grain already present in the matrix ('preformed nucleus'), that is capable of growth by virtue of its size and its orientation[130].† The grain structure and orientation, as set up in primary recrystallization and in subsequent grain growth, must be responsible for the coarsening behaviour on further annealing.

The interpretation of the incubation period in coarsening depends on the type of coarsening considered. In orientation-dependent coarsening (e.g. in 'cube texture' copper), where general grain growth is prevented by the low mobility of most of the matrix grain boundaries, the incubation period must be due to the initially low rate of growth of the 'nuclei' when they are still quite small. However, as shown by Rosi, Alexander, and Dube[133], an increase of the rate of growth with the size of the

* Following a calculation by E. Orowan (private communication), the necessary activation energy would be several orders of magnitude larger than that available at ordinary annealing temperatures.

† This orientation effect will be discussed in the following sections.

growing grain is encountered also in cases where general grain growth is prevented by second phase inhibition (' inhibition-dependent coarsening ').

In addition to the initially low rate of growth when the ' nuclei ' are still small, in the case of inhibition-dependent coarsening the incubation period may be also affected by the rate of re-solution and coalescence of the dispersed inhibiting particles. When the inhibition is initially very strong, as in the experiments with Al–Mn alloys[118], the decrease of inhibition during annealing is undoubtedly the predominant factor in determining the incubation period. The local conditions of decrease of inhibition play here an important role in the selection of sites for ' nucleation ' by permitting certain large matrix grains to grow in preference to others that may be equally large. The picture of ' nuclei ' for coarsening given above requires, therefore, some correction. In general, the ' nucleus ' for coarsening is a grain already present in the matrix, selected for growth by virtue of its large size, its orientation, and the relatively weak inhibition at its boundaries. Of these three selective factors the large relative size is always required, the importance of orientation increases with increasing sharpness of the matrix texture and that of local conditions of inhibition with the general level of inhibition in the material.

4.5. *Problems of Growth Rate*

(a) Theory of Boundary Migration

A rather detailed discussion of the problems of boundary migration was given in the recent review by Burke and Turnbull[10]. Consequently, the following comments are limited to a very brief summary. Anderson and Mehl's quantitative data[120] for the rate of growth of recrystallized grains in a fine grained, slightly deformed matrix of high purity Al, as a function of annealing temperature, lend themselves very well for theoretical analysis. It was pointed out by Mott[150] that the measured rates of growth are much too high for a theory in which the thermally activated elementary process is the transfer of a single atom from a crystal to the boundary layer. He found, however, that the difficulty can be eliminated if one assumes that the thermally activated elementary process is the transfer of a group of about twenty atoms into the boundary zone (' melting '). This theory accounts also for the surprisingly high activation energy for growth observed by Anderson and Mehl.

Turnbull[134] and Burke and Turnbull[10] treated the problem from the point of view of the absolute rate theory and they proposed an alternative interpretation, where the observed large variation of the rate of growth with temperature is supposed to be the result of progressive re-solution with increasing annealing temperature of growth-inhibiting dispersed particles. As discussed in § 4.3 of this chapter, the results of H. Suzuki[123] suggest that the metal used in Anderson and Mehl's investigations may in fact have contained some dispersion of a second phase. The apparent dependence of the measured activation energy for growth on the impurity content was noted by Suzuki. No doubt this correlation could be further

clarified if the impurities forming dispersed particles would be taken into account separately from those in solid solution. That the presence of second phase forming impurities does in fact increase the measured activation energy for the rate of boundary migration is seen when comparing the activation energy for grain growth in high purity brass (approximately 40 kcal/mole[108, 112]) with that in brass of commercial purity (approximately 60 kcal/mole[151, 152]). These observations tend to support Burke and Turnbull's point of view. On the other hand, since the effect of inhibition on the rate of boundary migration in recrystallization may be expected to be much larger if the preceding deformation is small than if it is large, the 'inclusion hypothesis' of Burke and Turnbull would require a great decrease in the measured apparent activation energy with increasing deformation. However, as noted by Burke and Turnbull themselves, the decrease of activation energy for the rate of growth actually observed by Anderson and Mehl is relatively small (fig. 47), in disagreement with the theory. Whether or not the unexpectedly high entropies of activation for growth calculated from Anderson and Mehl's data[134, 10] are in agreement with the theory can not be determined on the basis of available data. It now appears that a fully satisfactory solution of these problems will not be possible without additional data. What is needed is accurate growth rate measurements in a matrix, which is itself stable against structural changes during the growth anneal, and also free from dispersed phases.

(*b*) *Orientation Dependence of Boundary Mobility*

Dunn[155] measured the rate of growth of silicon–iron crystals as a function of their orientation relative to a fine grained matrix with essentially single orientation texture. He found a broad maximum of the rate of growth for a relative orientation corresponding to a [100] rotation of 45°. Independent observations by Lacombe and Berghezan[153] with high purity aluminium and by Tiedema, May and Burgers[154] with aluminium of commercial purity indicate that in recrystallization after small strains and in coarsening the large new grains do not absorb all matrix grains at the same rate, but that matrix grains with orientations near that of the large growing grain, or nearly twin-related to it, may persist over long periods of annealing, even when completely enveloped by the new grain. Since the grain boundaries corresponding to these two relative orientations have low interface energies[139], in the case of coarsening in commercial aluminium the above observation might be explained by assuming that the driving energy for the movement of these low energy boundaries is not sufficient to overcome the inhibition. However, Lacombe's results with recrystallization in high purity aluminium, where the inhibition, if any, is very much smaller, and at least a considerable fraction of the driving energy is due to the plastic strain, may be explained only on the assumption that the mobility of grain boundaries depends on the disorientation across them. Other observations[135] also support the view that, in general, low energy boundaries have low mobility. However, on the

basis of Guinier's results[41], it has been suggested[135] that low energy boundaries of a special type (essentially dislocation wall type subboundaries) may have high mobility in spite of their low energy. That large differences in mobility may exist between low energy boundaries of this special type and those of a more general kind is quite plausible, since the atomic mechanism of interface migration is different in the two cases.

It has been shown by Rathenau and Custers[156], Kronberg and Wilson[149] and Beck and Hu[157] that in face centred cubic metals the fastest growing grains on coarsening after random artificial nucleation in a polycrystalline matrix with single orientation texture are those with an orientation relative to the matrix corresponding to a 30° rotation around a [111] axis for Ni–Fe alloy and Cu, and to a 40° rotation around the same axis for Al (fig. 61, Plate 35). Beck, Sperry and Hu[55] later found the same orientation relationship for maximum boundary mobility in the recrystallization of slightly deformed high purity aluminium crystals artificially nucleated by scratching (fig. 62, Plate 35). Although additional work would be certainly very desirable to determine more quantitatively the orientation dependence of boundary mobility, the results of the above investigations are rather definite in establishing the orientation relationship for maximum boundary mobility in various f.c.c. metals.

This orientation (corresponding to 30° or 40° rotation around [111], which is approximately half way between the two relative orientations with very low interface energy mentioned above (corresponding to 0° and 60° rotations around [111], respectively), may be expected to correspond to a maximum of the interface energy. That there may be, at least qualitatively, a correlation between boundary mobility and interface energy, has been suggested[135] on the assumption that the rate determining process in boundary migration is akin to self diffusion along the boundary, and that the rate of boundary self diffusion increases with the boundary energy. Experimental confirmation has since been obtained[159] of the proposition that the rate of self diffusion is higher in high energy boundaries than in those of low energy. Turnbull's calculation[134], showing that in silver the free energy of activation for boundary migration is in fair agreement with that for grain boundary self diffusion, suggests that the elementary process in boundary migration may indeed be the same as in grain boundary self diffusion. More quantitative data would be desirable in order to test the above correlation.

Another attempt to rationalize the orientation dependence of boundary mobility might be based on the findings of Kronberg and Wilson[149], who observed that for a series of discrete rotation angles around a [111] axis and also around a [100] axis, in f.c.c. structures the atomic movements necessary for transition between the two disoriented lattices are particularly short and, therefore, presumably easy to accomplish, and approximately 14% of the atomic sites actually coincide, requiring no movement at all. Consequently, these discrete orientations might be assumed to correspond to maximum boundary mobility. Although two of the predicted [111] rotations are fairly close to the observed

orientations of greatest boundary mobility for Cu, respectively Al, a serious difficulty with this theory lies in the fact that most of the predicted orientations of maximum boundary mobility have not been confirmed experimentally. Furthermore, there is apparently no reason to assume that the relative orientations given by Kronberg and Wilson are the only ones possessing the described geometric features. As a matter of fact, in the case of twin orientation as many as one third of all atoms in the two f.c.c. lattices fully coincide and all the other corresponding sites are also less than the distance of closest packing apart when projected on the (111) plane. In spite of this, as mentioned above, boundaries between twin-related lattices are notoriously immobile. It may be, therefore, concluded that the suggestion just discussed is not likely to prove satisfactory.

4.6. *Annealing Textures*

It was noted above that the migration of high angle grain boundaries (recrystallization) involves local reorientation, as a grain of a certain orientation is replaced by a new grain of a different orientation. The observation that metals annealed after plastic deformation generally show a pronounced preferred orientation and that this is usually different from the deformation texture, indicates both that 'nuclei' for the new grains are available in orientations corresponding to the annealing texture, and that these 'nuclei' have the capacity to grow at the expense of the matrix during annealing. A theory of annealing textures must account for the scarcity of new grains in other than the preferred orientations. This scarcity may be attributed either to a shortage of nuclei in the corresponding orientations (oriented nucleation theory), or to their low rate of growth[135, 155, 157] (oriented growth theory).

The assumption that sharp annealing textures result from the availability of nuclei only in specific orientations is, in general, extremely difficult to verify directly, since nuclei are by nature small, and can not be identified as such until after they have grown and were, therefore, subject to growth selectivity. An oriented nucleation theory of annealing textures might, however, still have merit, if it would allow the prediction of annealing textures on the basis of some general principle regulating nucleus orientations, and applicable to a large number of individual cases. So far, no such general principle has emerged. Burgers and Louwerse[160] attempted to formulate one by assuming that nuclei in plastically strained metals are formed in regions of 'local curvatures', which are disoriented with regard to the main part of the lattice by rotation around a [112] axis parallel to the active slip plane and perpendicular to the active slip direction. They were able to analyse recrystallization textures in compressed and rolled aluminium single crystals in terms of such rotations, at least approximately. However, this particular oriented nucleation principle certainly did not prove generally applicable. Burgers himself, in a later publication[161] found it necessary to introduce a new oriented nucleation principle in order

to explain the formation of the ' cube texture ' (orientation : (100) [001]) in f.c.c. metals. According to this new principle, those elements of the rolled structure which have the cube orientation are the first to ' polygonize ' (i.e. they sustain subgrain growth at the highest rate) and therefore they are the first to become capable of acting as nuclei. However, this principle is not generally applicable either, since it does not account for annealing textures other than the cube texture. Actually, rolled 70–30 brass[162], cross-rolled copper[163], rolled single crystals of aluminium[160, 157] and of copper[165] etc. are known to recrystallize in other orientations. Even in straight-rolled copper, the amount of cube texture formed on annealing becomes appreciable only after very high deformation[164]. A third oriented nucleation principle was introduced by Kronberg and Wilson[149] to account for the texture obtained by coarsening in cube texture copper. It is apparent that, in the absence of a generally applicable oriented nucleation principle, in various specific instances new *ad hoc* assumptions are required regarding the orientation of the nuclei in order to explain the experimentally found annealing textures. So far, the oriented nucleation theory apparently did not allow the prediction of an unknown annealing texture.

The three specific oriented nucleation principles mentioned are subject to grave objections. The [112] rotations predicted by Burgers and Louwerse's principle were not confirmed by other investigators either for compressed[135] or rolled aluminium single crystals[157] or for rolled copper single crystals[165]. Furthermore, it was shown[157] that the results of Burgers and Louwerse themselves may be described more accurately by [111] rotations. The principle of Burgers, Liu and Tiedema predicts the formation of the cube texture on annealing, of deformed f.c.c. metals, certainly at least when a measurable fraction of the deformed material is in the cube orientation[161]. The investigation of Merlini and Beck[163] showed that in cross-rolled copper there is present a measurable fraction of material in the cube orientation, but the cube texture nevertheless does not form on annealing. This condition obtains in spite of the fact that the small fraction of cube oriented material in straight-rolled copper, where the cube texture does form on annealing, is only approximately half of that in cross-rolled copper. The principle of oriented nucleation advanced by Kronberg and Wilson[149] to explain the orientation of grains resulting from coarsening in cube texture copper predicts a large number of orientation relationships, of which actually only three are near orientations experimentally found. No suggestion was so far advanced to account for the absence of coarse grains near the other predicted orientations.

On the other hand, the oriented growth theory is based on the experimentally observable orientation relationship that corresponds to maximum boundary mobility. As stated in the previous section, for f.c.c. metals this orientation relationship may be described as a rotation of 30° or 40° around a [111] axis. By analogy, for h.c.p. metals maximum boundary mobility may be expected for a disorientation of 30° around

the [0001] axis[166]. When the deformed metal has a fairly sharp single orientation texture, the oriented growth theory predicts the formation of an annealing texture consisting of all the crystallographically equivalent orientation corresponding to maximum boundary mobility. In the case of copper, there are 8 equivalent orientations related to the ideal orientation of the deformation texture according to 30° rotations around the four [111] axes. Liu and Hibbard[165] found that, when a copper single crystal is heavily rolled on a (110) plane, in a [112] direction, and a sharp single orientation deformation texture is obtained, the resulting annealing texture in fact consists of the eight orientations mentioned above. The same eight relative orientations predominate on coarsening in f.c.c. Ni–Fe[156] and in Cu[168] with a single orientation cube texture matrix. In the case of hexagonal close packed metals, the expected orientation relationship corresponding to maximum boundary mobility is satisfied by one orientation only. Recent work of McGeary and Lustman[89] showed that a single orientation deformation texture may be obtained with polycrystalline zirconium by 60°-cross rolling. They also found that the annealing texure developing from this deformation texture corresponds to a rotation of 30° around the [0001] axis, as predicted on the basis of the expected maximum boundary mobility[169]. McHargue *et al.*[170] recently obtained an essentially single orientation deformation texture by straight-rolling a hexagonal close packed alloy of Ti+3·8% Al, and they found that the annealing texture, here too, was the one just described[171].

When the deformation texture is more complex, consisting of more than one orientation and of significant orientation spreads, the annealing texture corresponding to maximum growth rate in such a complex matrix is, in general, difficult to predict uniquely, without a detailed knowledge of the orientation dependence of boundary mobility. On the basis of the oriented growth theory, one may expect that the crystal orientations most favourable for growth in a complex matrix, and therefore actually developing in the corresponding annealing texture, will not, in general, necessarily have exactly the ideal orientation relationship with any one of the matrix texture components, but rather they may deviate to some extent because of the presence of other matrix texture components. Furthermore, several of the orientations ideally related to certain individual matrix texture components may be so unfavourably oriented for growth at the expense of other matrix components or of important matrix orientation spreads that they do not grow to a significant component in the annealing texture[163].

Despite the conditions described, in most of the known annealing textures of f.c.c. metals the various components of the annealing texture come surprisingly near to having the ideal orientation relationship with respect to at least one of the matrix texture components. In cross-rolled Cu[163] each of the eight main components of the annealing texture (type : (519)[$\bar{4}$, 11, 1]) is exactly so related to one of the two main

deformation texture components (type : (110) [$2\bar{2}3$]). The other eight potential annealing texture components, which actually do not form, are distinctly less favourably oriented* in relation to the second deformation texture component than the eight components that do form (table 1). In straight rolled 70–30 brass, the formation of the (112) [$1\bar{1}0$] annealing texture, which would be ideally oriented with respect to both main deformation texture components (type : (110) [$1\bar{1}2$]), is apparently prevented because it is very unfavourably oriented with respect to the principal orientation spread of the deformation texture (table 2). Instead, four main components develop on annealing, each one of which

Table 1. Cross-rolled Copper

Orientation relationship with second main def. text. comp't. (110) [$5\bar{5}7$]	Angular distance of rot. axis from nearest [111]	Angular distance of rot. axis from second nearest [111]	Angle of rotation
For actual recryst. text. components, related to first main def. texture comp. by 30° rotation around [111] near centre, (519) [$\bar{4}$, 11, 1]	16°	54·5°	49°
For hypothetical recryst. text. comp.'s, related to first main def. texture comp't. by 30° rotation around [111] at periphery	33·5°	39°	35°

is fairly close to having the ideal orientation relationship with one of the matrix components[145] and it is not very unfavourably oriented in relation to the other one and to the principal orientation spread (table 2). Also, in straight-rolled h.c.p. zirconium, which has a two-component deformation texture, the recrystallization texture[89, 167] corresponds quite well to a rotation around the [0001] axis of each of the deformation texture components and a deviation from ideal reorientation appears only as a slight difference in the angle of rotation (20° instead of 30°). It may be concluded, therefore, that the principal effect, if any, of the

* The angular distance of the axis of the rotation, which relates the two orientations, from the nearest [111] axis may be taken in a first, rough approximation as an index of the extent to which a certain orientation approximates that most favourable for growth (except near the identical and the twin orientations).

presence of additional texture components or major orientation spreads in the matrix is to prevent the formation of certain annealing texture components, which are unfavourably oriented for growth at their expense. The deviations from the ideal orientation relationship with one of the matrix texture components (as a result of the presence of other matrix

Table 2. Rolled 70–30 Brass

	Orientation relationship with		
	first main def. comp't. (111) [$1\bar{1}2$]		
	Angular distance of rotation axis from		Angle of Rotation
	nearest [111]	second nearest [111]	
For actual recryst. texture (225) [$73\bar{4}$]	11·5°	63·5°	32°
For hypothetical recryst. texture (112) [110]	0°	70·5°	30°
	second main def. comp't. (110) [$\bar{1}12$]		
For actual recryst. texture (225) [$73\bar{4}$]	15·5°	56·5°	46·7°
For hypothetical recryst. texture (112) [110]	0°	70·5°	30°
	centre of principal def. text. spread (110) [001]		
For actual recryst. texture (225) [$73\bar{4}$]	19°		45°
For hypothetical recryst. texture (112) [110]	43·5°		46°

orientations) are usually fairly small. This being so, at least approximate predictions usually can be made. The prediction made in the case of cross rolled copper prior to the quantitative determination of the annealing texture proved to be precise.

The same conclusions apply also for the formation of the cube texture ((100), [001]) in rolled f.c.c. metals. This annealing texture deserves particular attention, because of its unique sharpness, and because of the

rather detailed knowledge now available concerning its formation. It is known[172] that in copper the percentage cube texture formed increases with decreasing penultimate grain size and with increasing rolling reduction, in such a way that various combinations of these two variables leading to the same rolled grain thickness give rise to approximately the same fraction of cube oriented material, if the annealing treatment remains the same. It is also known that the fraction of the cube texture component increases with increasing annealing temperature[173] or time[174]. Finally, it has been observed[175] that the formation of the cube texture can be suppressed to a great extent if the rolled Ni–Fe alloy strip is etched to a thickness of approximately 1 micron before annealing. It was pointed out[176, 163] that not only the orientation of the cube texture and its sharpness but also the effect of all the above mentioned factors

Table 3. Cube Texture

	Orientation relationship with			
	Cu, main def. text. (5, 8, 14) [$\bar{1}\bar{1}$, $\bar{5}$, 7]*	Cu, centre of spread (112) [11$\bar{1}$]	brass, main def. text. (110) [1$\bar{1}$2]	brass centre of spread (110) [001]
Angular distance of rot. axis from nearest [111]	3°	23°	22°	55°
Angular distance of rot. axis from second nearest [111]	69°	50°	50°	55°
Angle of Rotation	45°	57°	56°	45°

* This 'ideal orientation' is more accurate than the (3, 5, 8) [3, $\bar{5}$, 2] orientation originally given by Hu, Sperry and Beck (see also ref. 186).

on its formation may be readily explained on the basis of the oriented growth theory. The cube orientation is not far removed from the ideal orientation relationship for maximum boundary mobility with the main deformation texture components in the rolled f.c.c. metals in which it does form on annealing (table 3), and it is in fact unique in having such an orientation relationship with all four deformation texture components[145]. For this reason, the growth of recrystallized grains in the cube orientation is favoured over others after the new grains have grown large enough to be in effective contact with matrix elements (e.g. grains, kinkbands) in more than one and, preferably, in all four matrix orientations. This critical size is reached relatively earlier and, therefore, selectively favoured growth of the cube grains may take place over a larger fraction of the total annealing period, if the elements of the rolled matrix are smaller

(hence the effect of penultimate grain size and of the rolling reduction) and if the annealing temperature is higher, or the annealing period longer. On the other hand, factors inhibiting the growth of the new grains (certain impurities) or preventing them from coming into effective contact with more matrix texture components (etching to small thickness) will tend to suppress the formation of the cube texture.

An objection raised[177] against the oriented growth theory was based on the observation that the maximum of the grain boundary energy, and presumably also of boundary mobility, as a function of disorientation is usually quite broad. Grains deviating from the orientation for maximum mobility by 15°, or so, still grow at relatively high rates. In view of such observations, it has been doubted that the orientation dependence of boundary mobility is sufficient to explain the sharp annealing textures known to occur. It has been noted[178], however, that the maximum becomes sharper if the increase of volume of recrystallized material rather than the linear rate of boundary migration is considered as a function of orientation; on the basis of volume, the pertinent variable is the third power of the mobility. If it is assumed that the 'incubation period' for nucleation is the period of slow 'invisible growth' and that its rate depends on the orientation relative to the matrix in a manner similar to that of visible growth, then the apparent rate of nucleation will be also similarly orientation dependent. Assuming that both G and N are proportional to the boundary mobility, the latter enters approximately at the fourth power into the exponential term of the formula giving the recrystallized volume. This is seen particularly well in the simplified form of the Johnson–Mehl equation (9),

$$F(t)=1-\exp\left(-\frac{\pi}{3}N\,.\,G^3t^4\right) \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where G and N are assumed to be independent of time. It is clear that the orientation dependence of the recrystallized volume may be relatively sharp, even if that of the mobility is not. An other effect, which further exaggerates the consequences of small differences in mobility, is due to the impingement on each other of the recrystallized grains[178]. As seen in fig. 62, fast growing grains 'cut in front' of grains with boundaries of lower mobility, so that the latter are 'pinched off'. As a result of this 'pinching off' due to impingement, grains with boundaries of maximum mobility can grow to disproportionately larger sizes than other grains, even if the differences in mobility are slight, provided that the number of competing new grains per unit volume is large enough to make the competition keen.* An additional reason, on the basis of oriented

* Burgers and Tiedema[129] proposed an explanation of the varying grain sizes obtained in this experiment on the basis of an assumed variation in the distance of the 'nuclei' from the edge of the specimen. It is clear, however, that this explanation does not account for the most significant effect, namely the observed correlation between size and orientation of the new grains.

growth, for the formation of sharp annealing textures comes into play if the matrix texture has more than one component. As already mentioned, the favoured orientations in this case are those for which the effective overall mobility in the complex matrix as a whole is largest. Such a ' compromise texture ' (e.g. the cube texture) may correspond to high mobility effective only in a very narrow range of orientations, because of the partially overlapping restrictions on mobility imposed by the various individual matrix texture components[166]. In view of these considerations, the objection raised appears to be ineffective.

An other objection raised against the oriented growth theory[177, 129] was that it is incompatible with the observed effect of annealing temperature on the resulting texture. Actually, the information now available on this effect is rather limited. It is known that in rolled Cu the fraction of cube texture increases with the annealing temperature[173] and that in cold rolled 70–30 brass annealing at 650°C results in a (110) [1$\bar{1}$2] type texture, in contrast to the (22$\bar{5}$) [734] texture obtained at around 340°C. In both of these cases the textural changes may be interpreted as the results of grain growth after recrystallization, the reorientation being in good agreement with the principles of oriented growth[145, 157]. A textural change, similar to that in brass, occurs also in silver, where it is associated with discontinuous grain growth[133]. In such cases where the annealing texture changes with the annealing temperature, it is possible to test whether or not the texture produced at the lower annealing temperature in fact changes on further annealing at the higher temperature into the same texture that is characteristic of the high annealing temperature. If such experiments would give positive results, they would confirm that the texture change is a result of additional growth. If the results turned out to be negative, explanation might be attempted for instance along one of the following lines: (1) Precipitation of dispersed particles of a second phase takes place during the low temperature anneal, inhibiting further growth, at least initially, on subsequent annealing at higher temperature, (2) the orientation dependence of boundary mobility changes with tenperature, or (3) the orientation dependence of nucleation rates changes with annealing temperature. If the incubation period is considered as the period of slow ' invisible growth ', then the effect on the rates of nucleation also may be a consequence of the change in the orientation dependence of boundary mobility. New data will be certainly necessary in order to allow a choice between the various possibilities.

A specific limitation of the oriented growth theory of annealing textures should be clearly noted. As seen above, well defined annealing textures are predicted only under circumstances where a reasonably large number of well disoriented nuclei are available. It was shown by various investigators[179, 180, 181] that, in the case of slightly deformed single crystals, where nuclei, and particularly well disoriented ones, are rather scarce the orientation of the recrystallized grains bears no simple relationship

with the matrix orientation. Apparently, the various nuclei are so far apart that competition between them is ineffective ; almost any nucleus can grow, with the exception of those in nearly identical or nearly twin orientations. This condition may be so exaggerated that active nuclei in favourable orientations are practically absent. For instance, carefully prepared specimens of cube texture copper may be annealed for long periods at 1000°c without coarsening[136], since the texture is so perfect that grains disoriented more than a few degrees are practically absent. If the specimen is annealed long enough, coarsening eventually does occur and the large grains have orientations in the order of 10° removed from that of the matrix. Kronberg and Wilson[149] and Dunn and Sharp[182] found some coarse grains rotated about 20° approximately around ⟨100⟩ axes with respect to the cube texture matrix in Cu, even when ⟨111⟩-rotated coarse grains also formed. It may be assumed that in borderline cases, where nuclei in orientations most favourable for growth are still fairly scarce, coarse grains also are able to develop in somewhat less favourable orientations. From the observed orientations one may conclude that, among the orientations differing from the matrix by up to about 20°, those corresponding approximately to ⟨100⟩-rotations are *relatively* most favourable for growth. Various investigators[167, 149, 135] noted that in cases of scarce nucleation the artificial introduction of sufficiently random nuclei gives rise to the growth of grains in the usual orientations corresponding to maximum boundary mobility. This furnishes additional proof that the special conditions discussed above are indeed characterized by a scarcity of nuclei. A different situation where nuclei are also absent, even though the matrix is capable of supporting the growth of ' nuclei ' artificially introduced, may arise in cases of inhibition by a dispersion. As described in § 4.3 (*c*), a larger grain may be able to grow in such a matrix at levels of inhibition and at annealing temperatures where the boundaries between the small matrix grains are effectively immobilized. If a single large grain is produced by local overheating, under properly chosen conditions this may be made to grow and to absorb the whole specimen into a single crystal. The method was used by Tiedema[183] and by Dunn and Lionetti[184] for producing crystals of predetermined orientations. Clearly, then, under carefully controlled conditions it is possible to produce in various ways orientations that are not determined by selective growth.

In a study of coarsening in rolled silicon–iron single crystals, Dunn[185] traced the orientation of the coarse grains to four larger than average size grains present in the structure at a stage of annealing before visible coarsening. He stated, although published no supporting evidence, that these particular grains grew to their large size at the expense of the deformed single crystal (primary recrystallization) rather than of the surrounding recrystallized grains. Dunn also stated that the orientation of these four grains was representative of that of all large grains in the structure deviating from the main orientation of the matrix. Thus, in

his view, the grains in question served as oriented nuclei for coarsening, without regard to the dependence of boundary mobility on the orientation of the growing grains relative to the recrystallized matrix. If his view is accepted, this case may be considered as somewhat similar to the special situations mentioned in the previous paragraph. The cause of the large size of recrystallized grains of the particular orientation in question was not explained. More recently, Dunn[186] attempted to generalize this mechanism to other cases of coarsening. However, there appears to be no experimental evidence, nor any evident reason for the assumption that in all these cases the structure obtained by primary recrystallization contains sufficiently large grains only in the orientations predominating in coarsening. For example, unless an additional selective principle is introduced in the case of cube texture Cu, one might expect that the minor texture components assumed by Dunn to furnish nuclei for coarsening would occur with equal frequency an all the 28 orientations which, together with the cube texture itself, may be derived from the four rolling texture components. (If one prefers to use the 22° and 38° [111]-rotations proposed by Kronberg and Wilson, as does Dunn, the number of differently oriented minor textures expected is much larger yet.) Since the coarsening texture actually consists of only four[149] or eight[168] components, it appears doubtful that Dunn's theory is capable of predicting, or even explaining, actual coarsening textures in face centred cubic metals, without additional *ad hoc* hypothesis.

As a summary, it may be stated that the only situations so far definitely established, where the orientation of grains formed on annealing is not determined by selective growth, but presumably by the available individual nucleus orientations, are the special ones characterized by scarcity of nuclei in orientations favourable for growth. In these cases the oriented growth theory does not apply, since the prerequisites for selective growth are not fulfilled, there being not enough nuclei to select from. It is interesting that even in slightly deformed single crystals the orientations of the naturally available nuclei are apparently quite random since, in the absence of selective growth, no well defined annealing textures are formed. On the other hand, after severe deformation, where one might expect much more random nucleation, since the local disorientations are much greater, in fact sharp annealing textures are formed. This may be ascribed most likely to selective growth, made possible by the abundance of nuclei. The observed predominance of the same kind of reorientation, corresponding to [111]-rotations of 30° or 40°, in both primary recrystallization and coarsening in all f.c.c. metals naturally follows from the oriented growth theory and from the observation that this particular orientation relationship also corresponds to maximum boundary mobility. In contrast to this, no oriented nucleation principle so far published proved to be suitable for generalization, i.e. for explaining more than the one case for which it was originally invented. The close similarity or identity of the predominant orientation relationship in a

number of carefully documented cases of both coarsening and primary recrystallization in f.c.c. metals with that for maximum boundary mobility would seem to require of a more general oriented nucleation theory that it assume for the nuclei in both primary recrystallization and coarsening the same orientations that are known to correspond to maximum boundary mobility. An ' oriented nucleation ' theory of this kind, as proposed recently by Liu and Hibbard[187], is somewhat difficult to distinguish from the oriented growth theory, except insofar as it is based on hypothetical, and not independently verifiable nucleus orientations, instead of on the experimentally observed orientation dependence of boundary mobility. In the case of a matrix with a sharp single orientation texture it leads to the same predictions as the oriented growth scheme, but it can not account without additional *ad hoc* hypotheses for the absence of some of the predicted reorientations in the case of complex matrix textures. Also, such an oriented nucleation theory would have great difficulty in explaining the extreme sharpness of the cube texture[145] formed on annealing from a rolling texture that includes very considerable orientation spreads. If it is assumed that the nuclei have a definite crystallographic relation to the matrix, one would expect each nucleus to be so related with respect to that microscopic portion of the matrix directly surrounding it ; consequently, the spread of nucleus orientations should correspond approximately to the spread of matrix orientations. In the theory of Liu and Hibbard[187] it is assumed that on recrystallization of rolled Cu nuclei are contributed mainly by matrix elements having the ' modified (358) [$3\bar{5}2$] ' ideal orientations of the rolling texture, and to a very minor extent by matrix elements with the (110) [$1\bar{1}2$] orientation. The continuous spread of matrix orientations, for instance that between the two orientations mentioned, is arbitrarily assumed to contribute no nuclei.

Acknowledgments

The author wishes to express his appreciation to Professors E. R. Parker and J. E. Dorn, Drs. J. Washburn, R. R. Eggleston, P. Gordon, H. J. Lambot and A. Kelly for kindly furnishing him with original drawings and photographs resulting from their own investigations, to Dr. Lambot and Dr. Kelly also for reading parts of the manuscript, to Mr. A. Merlini for reading the section on annealing textures, and to Dr. H. Hu and Messrs. J. Towers and P. R. Sperry for their excellent unpublished experimental contributions referred to in the paper. Their work was supported by the Office of Naval Research, U.S. Navy, under contract No. N6–ori–165, T.O.1. The preparation of this paper was partly supported by the U.S. Atomic Energy Commission, through contract No. AT (11–1)–67, Project No. 9.

References

(1) G. I. Taylor and H. Quinney, 1934, The Latent Energy Remaining in a Metal After Cold Working, *Proc. Roy. Soc.* A, **143,** 307.

(2) H. Suzuki, 1949, The Release of Energy Associated with Crystal Restoration Process in Cold-Worked Polycrystalline Copper, *Sci. Rep. RITU* A-1, 55.

(3) L. M. Clarebrough, M. E. Hargreaves, D. Michell and G. W. West, 1952, The Determination of the Energy Stored in a Metal During Plastic Deformation, *Proc. Roy. Soc.* A, **215,** 507.
(4) M. B. Bever and L. B. Ticknor, 1953, The Energy Stored During the Cold Working of a Gold–Silver Alloy, *Acta Met.*, **1,** 116.
(5) P. Gordon, personal communication.
(6) W. G. Burgers, 1941, Rekristallisation, Verformter Zustand und Erholung in Masing, *Handbuch der Metallphysik*, III.
(7) C. S. Barrett, 1953, *Structure of Metals*, 2nd Ed.
(8) A. Guinier, 1952, Substructures in Crystals, in Schockley : *Imperfections in Nearly Perfect Crystals* (Wiley), 402.
(9) J. E. Burke, 1948, The Fundamentals of Recrystallization and Grain Growth in *Grain Control in Industrial Metallurgy* (Cleveland : American Society for Metals), 1.
(10) J. E. Burke and D. Turnbull, 1952, Recrystallization and Grain Growth, in Chalmers : *Progress in Metal Physics*, III, 220.
(11) Paul A. Beck, 1952, Interface Migration in Recrystallization in *Metal Interfaces* (Cleveland : American Society for Metals), 208.
(12) C. Crussard, 1953, Generalites et Considerations Theoriques sur la Recristallisation. *Metaux, Corrosion, Industries*, **28.**
(13) P. Lacombe, 1953, Le Grossissement du Grain des Metaux. *Metaux, Corrosion, Industries*, **28,** 378.
(14) M. Polanyi and E. Schmid, 1923, *Verh. dtsch. phys. Ges., Berlin*, **4.** See also M. Polanyi and E. Schmid, 1925, Verfestigung und Entfestigung von Sn-Kristallen, *Z. Phys.*, **32,** 684.
(15) O. Haase and E. Schmid, 1925, Über den Gleitwiderstand von Metallkristallen, *Z. Phys.*, **33,** 413.
(16) C. H. Li, J. Washburn and E. R. Parker, 1953, Variation of Plastic Properties with Annealing Procedure in Zinc Single Crystals, *Trans. Amer. Inst. metall. Engrs*, **197,** 1223.
(17) G. Masing, 1944, Zur Aktivierungsenergie der Erholung von durch Kaltreckung Verfestigten Metallen, *Z. Metallk.*, **36,** 173.
(18) R. W. K. Honeycombe, 1951–52, Inhomogeneities in the Plastic Deformation of Metal Crystals I, *J. Inst. Met.*, **80,** 45.
(19) R. Drouard, J. Washburn and E. R. Parker, 1953, Recovery in Single Crystals of Zinc, *Trans. Amer. Inst. metall. Engrs*, **197,** 1226.
(20) W. A. Johnson and R. F. Mehl, 1939, Reaction Kinetics in Processes of Nucleation and Growth, *Trans. Amer. Inst. metall. Engrs*, **135,** 416.
(21) G. Masing and J. Raffelsieper, 1950, Mechanische Erholung von Aluminium-Einkristallen, *Z. Metallk.*, **41,** 65.
(22) D. Kuhlmann, G. Masing and J. Raffelsieper, 1949, Zur Theorie der Erholung, *Z. Metallk.*, **40,** 241.
(23) D. Kuhlmann, 1947, Zur Theorie der Nachwirkungserscheinungen, *Z. Phys.*, **124,** 468.
(24) P. H. Miller and R. F. Banks, 1942, Self Diffusion in Zinc, *Phys. Rev.*, **61,** 648.
(25) T. E. Tietz, R. A. Anderson and J. E. Dorn, 1949, Effect of Prestraining Temperatures on the Recovery of Cold Worked Aluminium, *Trans. Amer. Inst. metall. Engrs*, **185,** 921.
(26) G. Borelius, S. Berglund and S. Sjoberg, 1952, Measurements on the Evolution of Heat During the Recovery of Cold-Worked Metals, *Ark. Fys.*, **6,** 143.
(27) A. E. van Arkel and W. G. Burgers, 1928, Verbreiterung der Debye-Scherrerschen Linien von kaltbearbeitetem Wolframdraht und Wolframband als Funktion der Glühtemperatur und Glühdauer, *Z. Phys.*, **48,** 690.

(28) J. E. Wilson and L. Thomassen, 1934, X-ray Line Broadness of Metals and Alloys and its Relation to High Temperature Stability, *Trans. Amer. Soc. Metals*, **22,** 769.

(29) B. L. Averbach, 1949, Recovery and Recrystallization in Brass, *Trans. Amer. Inst. metall. Engrs*, **185,** 491.

(30) A. H. Lutts and P. A. Beck, 1954, Annealing of a Cold Rolled Aluminum Single Crystal, *Trans. Amer. Inst. metall. Engrs*, **200,** 257.

(31) T. V. Cerian, P. Pietrokowsky and J. E. Dorn, 1949, Some Observations on the Recovery of Cold Worked Aluminum, *Trans. Amer. Inst. metall. Engrs*, **185,** 948.

(32) M. S. Paterson and E. Orowan, 1948, X-ray Line Broadening in Cold-Worked Metals, *Nature, Lond.*, **162,** 991.

(33) G. K. Williamson and W. H. Hall, 1953, X-ray Line Broadening from Filed Aluminium and Wolfram, *Acta Met.*, **1,** 22.

(34) J. S. Smart, Jr., A. A. Smith, Jr. and A. J. Phillips, 1941, Preparation and Some Properties of High Purity Copper, *Trans. Amer. Inst. metall. Engrs*, **143,** 272.

(35) J. Molenaar and W. H. Aarts, 1950, Change of Resistivity by Cold Working at Liquid-Air Temperature, *Nature, Lond.*, **166,** 690.

(36) T. Broom, 1952, The Effect of Temperature of Deformation on the Electrical Resistivity of Cold-Worked Metals and Alloys, *Proc. Phys. Soc.* B, **65,** 871.

(37) R. R. Eggleston, 1952, Cold Work Studies on Copper at Low Temperatures, *J. Appl. Phys.*, **23,** 1400, and personal communication.

(38) G. Tammann, 1934, Die Erholung Metallischer Werkstoffe von den Folgen der Kaltbearbeitung, *Z. Metallk.*, **26,** 97.

(39) Henry S. Rawdon and Torkel Berglund, 1927–28, Unusual Features in the Microstructure of Ferrite, *Scientific Papers of the Bureau of Standards*, **22**, 649, Paper S-571.

(40) A. B. Greninger, 1936, Transformation Twinning of Alpha Iron, *Trans. Amer. Inst. metall. Engrs*, **120,** 293.

(41) A. Guinier and P. Lacombe, 1948, L'état " polygonise " du crystal metallique, *Métaux et Corros.*, **23,** 212.

(42) J. A. Collins and C. H. Mathewson, 1940, Plastic Deformation and Recrystallization of Aluminum Single Crystals, *Trans. Amer. Inst. metall. Engrs*, **137,** 150.

(43) C. Crussard, 1944, Etude de Recuit de l'Aluminium, *Rev. Metall.*, **41,** 118.

(44) P. Lacombe and L. Beaujard, 1948, Les imperfections de structure des cristeaux uniques d'aluminium pur, *Rev. Metall.*, **45,** 317.

(45) P. Lacombe, 1948, Sub-Boundary and Boundary Structures in High Purity Aluminium, *Bristol Conference Report* (London : Physical Society), 91.

(46) A. Guinier and J. Tennevin, 1951, Researches on the Polygonization of Metals, *Progress in Metal Physics* II (New York and London), 177.

(47) C. G. Dunn and F. W. Daniels, 1951, Formation and Behavior of Subboundaries in Silicon-Iron Crystals, *Trans. Amer. Inst. metall. Engrs*, **191,** 147.

(48) W. A. Wood and R. F. Scrutton, 1950, Mechanism of Primary Creep in Metals, *J. Inst. Met.*, **77,** 423.

(49) J. Hino, P. G. Shewmon and P. A. Beck, 1952, Effect of Simultaneous Strain on Subgrain Growth, *Trans. Amer. Inst. metall. Engrs*, **194,** 873.

(50) O. D. Sherby, A. Goldberg and J. E. Dorn, 1952, Effect on Prestrain Histories on the Creep and Tensile Properties of Aluminum, *21st Technical Report to ONR*, October.

(51) Footnote on p. 234 of ref. (11).
(52) J. Washburn and E. R. Parker, 1952, Kinking in Zinc Single-Crystal Tension Specimens, *Trans. Amer. Inst. metall. Engrs*, **194,** 1076.
(53) C. H. Li, E. H. Edwards, J. Washburn and E. R. Parker, 1953, Stress Induced Movement of Crystal Boundaries, *Acta Met.*, **1,** 223.
(54) Discussion of ref. (52) by P. A. Beck, 1953, *Trans. Amer. Inst. metall. Engrs*, **197,** 737.
(55) P. A. Beck, P. R. Sperry and H. Hu, 1950, The Orientation Dependence of the Rate of Boundary Migration, *J. appl. Phys.*, **21,** 420.
(56) J. M. Burgers, 1940, Geometrical Considerations Concerning the Structural Irregularities to be Assumed in a Crystal, *Proc. Phys. Soc.* (*London*), **52,** 23.
(57) Authors' answer to Discussion of Ref. (52), 1953, *Trans. Amer. Inst. metall. Engrs*, **197,** 737.
(58) D. W. Bainbridge, C. H. Li, E. H. Edwards, J. Washburn and E. R. Parker, 1954, Recent Observations on the Motion of Small Angle Dislocation Boundaries, *Acta Met.*, **2,** 322.
(59) W. Schockley and W. T. Read, 1950, Dislocation Models of Crystal Grain Boundaries, *Phys. Rev.*, **78,** 275.
(60) W. A. Wood and H. J. Tapsell, 1946, Mechanism of Creep in Metals, *Nature, Lond.*, **158,** 415.
(61) G. R. Wilms and W. A. Wood, 1949, Mechanism of Creep in Metals, *J. Inst. Met.*, **75,** 693.
(62) W. A. Wood and W. A. Rachinger, 1949, The Mechanism of Deformation in Metals, With Special Reference to Creep, *J. Inst. Met.*, **76,** 237.
(63) R. W. Cahn, 1949, Recrystallization of Single Crystals After Plastic Bending, *J. Inst. Met.*, **76,** 121.
(64) R. W. Cahn, 1951, Slip and Polygonization in Aluminium, *J. Inst. Met.*, **79,** 129.
(65) A. H. Cottrell, 1949, Theory of Dislocations, *Progress in Metal Physics* I, 77.
(66) G. B. Greenough and Edna M. Smith, 1950, The Mechanism of Creep as Revealed by X-ray Methods, *J. Inst. Met.*, **77,** 435.
(67) W. A. Wood, 1950, In discussion, *J. Inst. Met.*, **77,** 606.
(68) R. W. Cahn, 1950, Internal Strains and Recrystallization, *Progr. Metal Phys.* II, 151.
(69) A. Guinier and J. Tennevin, 1948, *Acta Cryst.*, **1,** 188.
(70) H. J. Lambot, L. Vassamillet, J. Dejace, 1935, Sur la mesure des desorientations reticulaires dans les monocristaux metallique, *Acta Met.*, **1,** 711 and personal communication.
(71) R. D. Heidenreich, 1951, Electron Transmission Through Thin Metal Sections, *Bell Syst. tech. J.*, October.
(72) J. N. Kellar, P. B. Hirsch and J. S. Thorp, 1950, An X-ray Micro-Beam Examination of a Plastically Deformed Metal, *Nature, Lond.*, **165,** 554.
(73) Paul A. Beck and Hsun Hu, 1952, Annealing Textures in Rolled Face-Centered Cubic Metals, *Trans. Amer. Inst. metall. Engrs*, **194,** 83.
(74) P. Gay and A. Kelly, 1953, X-ray Studies of Polycrystalline Metals Deformed by Rolling I, *Acta Cryst.*, **6,** 165.
(75) N. K. Chen and R. Maddin, 1953, Cold Rolling and Annealing Textures of Molybdenum Single Crystals, *Trans. Amer. Inst. metall. Engrs*. **197,** 300.
(76) G. P. Conard II, B. L. Averbach and M. Cohen, 1953, Plastic Bending of Zinc Crystals, *Trans. Amer. Inst. metall. Engrs*, **197,** 1036.

(77) P. B. Hirsch, 1952, A Study of Cold-Worked Aluminum by an X-ray Micro-Beam Technique III, *Acta Cryst.*, **5,** 172.
(78) T. Broom, 1954, Lattice Defects and Electrical Resistivity of Metals, *Advances in Physics*, **3,** 26.
(79) E. Orowan, 1942, A Type of Plastic Deformation New in Metals, *Nature, Lond.*, **149,** 643.
(80) D. C. Jillson, 1950, An Experimental Survey of Deformation and Annealing Processes in Zinc, *Trans. Amer. Inst. metall. Engrs*, **188,** 1009.
(81) J. B. Hess and C. S. Barrett, 1949, Structure and Nature of Kink Bands in Zinc, *Trans. Amer. Inst. metall. Engrs*, **185,** 599.
(82) G. Masing and M. Polanyi, 1923, Kaltreckung und Verfestigung, *Ergebn. Exakt. Naturw.*, II, 177.
(83) D. Kuhlmann-Wilsdorf and H. Wilsdorf, 1953, The Surface Structures of Deformed Aluminum, Copper, Silver and Alpha-Brass and Their Theoretical Interpretation, *Act. Met.*, **1,** 394.
(84) E. R. Parker and T. H. Hazlett, 1953, Principles of Solution Hardening, 12*th Technical Report to AEC* (October).
(85) A. H. Cottrell, 1949, Theory of Dislocations in Chalmers *Progress in Metal Physics*, I, 77.
(86) For a review see P. A. Beck, 1953, Notes on Work Hardening and Recovery, *Acta Met.*, **1,** 422.
(87) M. Semchyshen and G. A. Timmons, 1952, Preferred Orientation of Arc-Cast Molybdenum Sheet, *Trans. Amer. Inst. metall. Engrs*, **194,** 279.
(88) C. J. McHargue, personal communication.
(89) R. K. McGeary and B. Lustman, 1953, Kinetics of Thermal Reorientations in Cold Rolled Zirconium, *Trans. Amer. Inst. metall. Engrs*, **197,** 284.
(90) Discussion by C. G. Dunn, 1949, *Trans. Amer. Inst. metall. Engrs*, **185,** 309.
(91) C. S. Barrett and F. W. Steadman, 1942, Structure of Copper After Rolling, *Trans. Amer. Inst. metall. Engrs*, **147,** 57.
(92) R. F. Mehl, 1948, Recrystallization, *Metals Handb.* (American Society for Metals), 259.
(93) L. J. Demer and P. A. Beck, 1949, Effect of Composition on Grain Growth in Aluminum–Magnesium Solid Solutions, *Trans. Amer. Inst. metall. Engrs*, **180,** 147.
(94) H. J. Wallbaum and R. Mischer, 1949, Die Mikroharte rekristallisierender Aluminiumblaeche, *Z. Metallk.*, **40,** 179.
(95) Paul A. Beck, 1952, Do Metals Recrystallize ?, *Trans. Amer. Inst. metall. Engrs*, **194,** 979.
(96) N. F. Mott, 1953, A Theory of Work-Hardening of Metals II. *Phil. Mag.*, **44,** 742.
(97) H. Alterthum, 1922, Zur Theorie der Rekristallisation, *Z. Metallk.*, **14,** 417.
(98) A. E. van Arkel and M. G. van Bruggen, 1928, Rekristallisationserscheinungen bei Aluminium II, *Z. Phys.*, **51,** 520.
(99) L. W. Eastwood, A. E. Bousu and C. T. Eddy, 1935, Recrystallization and Grain Growth in Cold Worked Polycrystalline Metals, *Trans. Amer. Inst. metall. Engrs*, **117,** 246.
(100) E. E. Stansbury, G. E. Elder and M. L. Picklesimer, 1953, Calorimetric Studies of Plastic Deformation and Phase Transformations, *AEC Report ORO* 86, March.
(101) L. M. Clarebrough, M. E. Hargreaves and G. W. West, 1953, The Release of Energy Stored in Deformed Nickel, *Phil. Mag.*, **44,** 915.

(102) C. S. Smith, 1951, in discussion: *The Physics of Powder Metallurgy*, p. 50 (New York: McGraw-Hill Book Co.).
(103) J. A. Ewing and W. Rosenhain, 1900, Crystalline Structure of Metals, *Phil. Trans. Roy. Soc.* A, **193,** 353.
(104) G. Tammann, 1912, Über die Änderung d. Eigenschaften bei der Bearbeitung von Metallen, *Z. Elektrochem.*, **18,** 584.
(105) R. Vogel, 1923, Über den wechselseitigen Auf- und Abbau sich berührender metallischer Kristallite im Konglomerat, *Z. anorg. Chem.*, **126,** 1.
(106) D. Harker and E. R. Parker, 1945, Grain Shape and Grain Growth, *Trans. Amer. Soc. Metals*, **34,** 156.
(107) C. S. Smith, 1948, Grains, Phases and Interfaces: An Interpretation of Microstructure, *Trans. Amer. Inst. metall. Engrs*, **175,** 15.
(108) J. E. Burke, 1949, Some Factors Affecting the Rate of Grain Growth in Metals, *Trans. Amer. Inst. metall. Engrs*, **180,** 73.
(109) J. E. Burke and Y. G. Shiau, 1948, The Effect of Mechanical Deformation on Grain Growth in Alpha Brass, *Trans. Amer. Inst. metall. Engrs*, **175,** 141.
(110) C. S. Smith, 1949, In discussion, *Trans. Amer. Inst. metall. Engrs*, **185,** 112.
(111) P. A. Beck, 1948, Effect of Recrystallized Grain Size on Grain Growth, *J. appl. Phys.*, **19,** 507.
(112) Unpublished data by John Towers, Jr. and P. A. Beck, 1949.
(113) P. A. Beck, J. C. Kremer, L. J. Demer and M. L. Holzworth, 1948, Grain Growth in High Purity Aluminum and in an Aluminum–Magnesium Alloy, *Trans. Amer. Inst. metall. Engrs*, **175,** 372.
(114) P. A. Beck, M. L. Holzworth and Hsun Hu, 1948, Instantaneous Rates of Grain Growth, *Phys. Rev.*, **73,** 526.
(115) E. S. Machlin, 1953, Some Applications of the Thermodynamic Theory of Irreversible Processes to Physical Metallurgy, *Trans. Amer. Inst. metall. Engrs*, **197,** 437.
(116) J. E. Burke, 1951, The Migration of Grain Boundaries, in *Atom Movements*, p. 209 (Cleveland: American Society for Metals).
(117) R. L. Fullman, 1952, Boundary Migration During Grain Growth, in *Metal Interfaces*, p. 179 (Cleveland: American Society for Metals).
(118) P. A. Beck, M. L. Holzworth and P. R. Sperry, 1949, Effect of a Dispersed Phase on Grain Growth in Al–Mn Alloys, *Trans. Amer. Inst. metall. Engrs*, **180,** 163.
(119) J. K. Stanley and R. F. Mehl, 1942, Recrystallization of Silicon Ferrite in terms of Rate of Nucleation and Rate of Growth, *Trans. Amer. Inst. metall. Engrs*, **150,** 260.
(120) W. A. Anderson and R. F. Mehl, 1945, Recrystallization of Aluminum in terms of Rate of Nucleation and Rate of Growth, *Trans. Amer. Inst. metall. Engrs*, **161,** 140.
(121) W. A. Johnson and R. F. Mehl, 1939, Reaction Kinetics in Processes of Nucleation and Growth, *Trans. Amer. Inst. metall. Engrs*, **135,** 416.
(122) R. Karnop and G. Sachs, 1930, Zur Kinetik der Rekristallisation, *Z. Phys.*, **60,** 464.
(123) H. Suzuki, 1953, Growth Rates of New Grains during the Primary Recrystallization of Aluminium Plates, *Sci. Rep. RITU*, A **5,** 414.
(124) A. E. van Arkel and J. J. A. Ploos van Amstel, 1928, Rekristallisationserscheinungen beim Zinn, *Z. Phys.*, **51,** 534.
(125) H. Hu and P. A. Beck, 1948, Unpublished research.
(126) B. F. Decker and D. Harker, 1950, Activation Energy for Recrystallization in Rolled Copper, *Trans. Amer. Inst. metall. Engrs*, **188,** 887.

(127) W. E. Seymour and D. Harker, 1950, Recrystallization Reaction Kinetics and Texture Studies of a 50 Iron 50 Nickel Alloy, *Trans. Amer. Inst. metall. Engrs*, **188,** 1001.

(128) P. A. Beck and P. R. Sperry, 1949, Effect of Recrystallization Texture on Grain Growth, *Trans. Amer. Inst. metall. Engrs*, **185,** 240. See also: P. A. Beck in discussion, *Trans. Amer. Inst. metall. Engrs*, **185,** 311.

(129) W. G. Burgers and T. J. Tiedema, 1953, Notes on the Theory of Annealing Textures, *Acta Met.*, **1,** 234.

(130) J. E. Burke, 1949, In discussion, *Trans. Amer. Inst. metall. Engrs*, **185,** 315. See also authors' answer to discussion.

(131) W. Feitknecht, 1926, Crystal Growth in Recrystallized Cold Worked Metals, *J. Inst. Met.*, **35,** 131.

(132) R. Karnop and G. Sachs, 1929, Die Grobkristallisation von Aluminium, *Metallwirtschaft*, **8,** 1115.

(133) F. D. Rosi, B. H. Alexander and C. A. Dube, 1952, Kinetics and Orientation Relationships of Secondary Recrystallization in Silver, *Trans. Amer. Inst. metall. Engrs*, **194,** 189.

(134) D. Turnbull, 1951, Theory of Grain Boundary Migration Rates, *Trans. Amer. Inst. metall. Engrs*, **191,** 661.

(135) C. S. Barrett, 1940, Recrystallization Texture in Aluminum after Compression, *Trans. Amer. Inst. metall. Engrs*, **137,** 128.

(136) P. A. Beck and P. R. Sperry, 1951, Unpublished work.

(137) R. Becker, 1926, Über Plastizität, Verfestigung und Rekristallisation, *Z. tech. Phys.*, **7,** 547.

(138) A. H. Cottrell, 1953, Theory of Dislocations, in Chalmers' *Progress in Metal Physics*, IV, p. 205–264.

(139) K. T. Aust and B. Chalmers, 1952, Energies and Structure of Grain Boundaries, in *Metal Interfaces* (Cleveland: American Society for Metals), pp. 153–178.

(140) R. W. Cahn, 1950, A New Theory of Recrystallization Nuclei, *Proc. Phys. Soc.* A, **63,** 323.

(141) W. G. Burgers, 1947, Recovery and Recrystallization viewed as Processes of Dissolution and Movement of Dislocations, *Proc. Kon. Nederl. Akad. v. Wetenschappen*, **50,** 452.

(142) W. L. Bragg, 1940, In discussion, *Proc. Phys. Soc.* **52,** 54.

(143) W. G. Burgers, 1949, Crystal Growth in the Solid State, *Physica*, **15,** 92.

(144) P. A. Beck, 1949, The Formation of Recrystallization Nuclei, *J. appl. Phys.*, **20,** 633.

(145) P. A. Beck and H. Hu, 1952, Annealing Textures in Rolled Face-Centered Cubic Metals, *Trans. Amer. Inst. metall. Engrs*, **194,** 83.

(146) P. A. Beck and P. R. Sperry, 1950, Strain Induced Grain Boundary Migration in High Purity Aluminum, *J. appl. Phys.*, **21,** 150. (See also ref. 11.)

(147) A. Guinier and J. Tennevin, 1949, Comparison of the Perfection of Crystals of Primary and Secondary Recrystallization, *Philips Res. Rep.*, **4,** 316.

(148) T. J. Tiedema, 1950, X-ray Investigation of the Nuclear Spot of Crystals Obtained by Recrystallization, *Proc. Kon. Nederl. Akad. Wetenschappen*, **53,** 1422.

(149) M. L. Kronberg and F. H. Wilson, 1949, Secondary Recrystallization in Copper, *Trans. Amer. Inst. metall. Engrs*, **185,** 501.

(150) N. F. Mott, 1948, Slip at Grain Boundaries and Grain Growth in Metals, *Proc. Phys. Soc.*, **60,** 391.

(151) J. E. Burke, 1947, Grain Growth in Alpha-Brass, *J. appl. Phys.*, **18,** 1028.
(152) P. A. Beck, W. D. Manly and J. Towers, 1948, Grain Growth in 70-30 Brass, *Trans. Amer. Inst. metall. Engrs*, **175,** 162.
(153) P. Lacombe and A. Berghezan, 1949, Relation d'orientation entre monocrystaux metalliques de recrystallization et petits cristaux inclus, *Metaux et Corros.*, **24,** 1.
(154) T. J. Tiedema, W. May and W. G. Burgers, 1949, Inclusions in Aluminium Crystals, *Acta Cryst.*, **2,** 151.
(155) C. G. Dunn, 1949, Recrystallization Textures, in *Cold Working of Metals* (Cleveland : American Society of Metals), p. 113.
(156) G. W. Rathenau and J. F. H. Custers, 1949, Secondary recrystallization of face-centered Ni–Fe alloys, *Philips Res. Rep.*, **4,** 241.
(157) P. A. Beck and H. Hu, 1949, Recrystallization Texture and Coarsening Texture in High Purity Aluminum, *Trans. Amer. Inst. metall. Engrs*, **185,** 627.
(158) D. Turnbull, 1951, Grain Boundary and Surface Diffusion, in *Atom Movements* (Cleveland : American Society of Metals), 129–152.
(159) M. R. Achter and R. Smoluchowski, 1951, Diffusion in Grain Boundaries and Their Structure, *J. appl. Phys.*, **22,** 1260.
(160) W. G. Burgers and P. C. Louwerse, 1931, Über den Zusammenhang zwischen Deformationsvorgang und Rekristallisationstextur bei Aluminium, *Z. Phys.*, **67,** 605.
(161) W. G. Burgers, Y. H. Liu and T. J. Tiedema, 1951, The Influence of Crystal Orientation on Polygonization of Aluminum Single Crystals, *Proc. Koninkl. Nederl. Acad. Wetenschappen*, **94,** 459.
(162) Goler and G. Sachs, 1929, Walz- und Rekristallisationstextur regularflachenzentrierter Metalle IV., *Z. Phys.*, **56,** 485.
(163) A. Merlini and P. A. Beck, 1953, Study of the Origin of the Cube Texture, *Acta Met.*, **1,** 598.
(164) W. M. Baldwin, Jr., 1946, Effect of Rolling and Annealing upon the Crystallography, Metallography and Physical Properties of Copper Strip, *Trans. Amer. Inst. metall. Engrs*, **166,** 591.
(165) Y. C. Liu and W. R. Hibbard, Jr., 1953, Recrystallization of a Cold-Rolled Copper Single Crystal, *Trans. Amer. Inst. metall. Engrs*, **197,** 672.
(166) P. A. Beck, 1951, Theory of Annealing Textures, *Trans. Amer. Inst. metall. Engrs.*, **191,** 475.
(167) J. H. Keeler, W. R. Hibbard, Jr., and B. F. Decker, 1953, Textures of Rolled and Annealed Iodide Zirconium, *Trans. Amer. Inst. metall. Engrs*, **197,** 932.
(168) J. S. Bowles and W. Boas, 1948, The Effect of Crystal Arrangement on " Secondary Recrystallization " in Metals, *J. Inst. Met.*, **24,** 501.
(169) R. K. McGeary and B. Lustman, personal communication.
(170) C. J. McHargue, Sam E. Adair, Jr. and J. P. Hammond, 1953, Effects of Solid Solution Alloying on the Cold-Rolled Texture of Titanium, *Trans. Amer. Inst. metall. Engrs*, **197,** 1199.
(171) C. J. McHargue, personal communication.
(172) M. Cook and T. Ll. Richards, 1940, The Structural Changes in Copper Effected by Cold Rolling and Annealing, *J. Inst. Met.*, **66,** 1.
(173) O. Dahl and F. Pawlek, 1936, Kornordnung und Kornwachstum bei Walzblechen, *Z. Metallk.*, **28,** 266.
(174) P. A. Beck, 1951, Orientation in Recrystallization and Grain Growth. Chapter 3 in *The Physics of Powder Metallurgy* (New York : McGraw-Hill Book Co.).
(175) J. F. H. Custers and G. W. Rathenau, 1941, Recrystallization in Rolled Nickel–Iron, *Physica*, **8,** 759.

(176) P. A. Beck, 1951, Origin of the Cube Texture in Face-Centered Cubic Metals, *Trans. Amer. Inst. metall. Engrs*, **191,** 474.

(177) W. G. Burgers, 1952, Recrystallization and Grain Growth in Solid Metals, in *Report on the 9th Solvay Conference*, L'Etat Solide, p. 431 (Bruxelles : R. Stoopes).

(178) P. A. Beck, 1953, Notes on the Theory of Annealing Textures, *Acta Met.*, **1,** 230.

(179) W. G. Burgers and J. C. M. Basart, 1929, Rekristallisation von Aluminiumeinkristallen, *Z. Phys.*, **54,** 74.

(180) A. Laloeuf and C. Crussard, 1951, Relations entre la Deformation et la Recristallisation de Monocristaux d'Aluminium, *Rev. Met.*, **48,** 461.

(181) N. K. Chen and C. H. Mathewson, 1952, Recrystallization of Aluminum Single Crystals after Plastic Extension, *Trans. Amer. Inst. metall. Engrs*, **194,** 501.

(182) C. G. Dunn and M. Sharp, 1952, Secondary Recrystallization Texture in Copper, *Trans. Amer. Inst. metall. Engrs*, **194,** 42.

(183) T. J. Tiedema, 1949, Preparation of Aluminium Single Crystals With a Definite Orientation, *Acta Cryst.*, **2,** 261.

(184) C. G. Dunn, 1949, Controlled Grain Growth Applied to the Problem of Grain Boundary Energy Measurements, *Trans. Amer. Inst. metall. Engrs*, **185,** 72.

(185) C. G. Dunn, 1953, Secondary Recrystallization Textures and Their Origin in Cold-Rolled Single Crystals of Silicon Iron, *Acta Met.*, **1,** 163.

(186) C. G. Dunn, 1954, On the Theory of Secondary Recrystallization Texture Formation in Face-Centered Cubic Metals, *Acta Met.*, **2**, 386.

(187) Y. C. Liu and W. R. Hibbard, Jr., On the Relationship of Texture Changes of Cold-Rolled Face-Centered Cubic Metals During Recrystallization (to be published).

Electrons in Lattice Fields

By H. Fröhlich
Department of Theoretical Physics, University of Liverpool

§ 1. Introduction

An electric point charge e brought into a dielectric medium polarizes its surroundings. If this point charge is static then the polarization is determined by the static dielectric constant. For a moving point charge the polarization can only be determined if the dynamic properties of the medium are known. If furthermore the point charge is replaced by a particle with a certain mass and charge, then the polarization produced by the particle reacts on it. A situation arises which can be described in terms of a field, the polarization field, which interacts with the particle; the latter then possesses a self energy in the field. A simple and interesting application of this idea is presented in the motion of a free electron in the conduction band of an ionic crystal with which the present article is mainly concerned.

The interest in discussing this question is twofold. Firstly, of course, a calculation of the properties of free electrons in ionic crystals has an intrinsic interest. Secondly, however, this case provides a very simple example for a non relativistic field theory, and in view of its simplicity it might be expected to lead to the discovery of a number of new features of such fields and to the development of new methods. In fact after the first application of the methods of field theory to electrons in ionic crystals (Fröhlich, Pelzer and Zienau 1950), the use of these methods in metals led to an important step in the theory of superconductivity and to the prediction of the isotope effect (Fröhlich 1950). Subsequent discussion has shown, however (Fröhlich 1953) that new methods are required to deal with all aspects of superconductivity, and at present it seems that a method which would combine the results of the two methods described in § 4 and § 5 might fit the requirements (cf. end of § 6).

The methods of § 4 and § 5 can best be described as dynamic and static respectively. The static method is essentially an application of Hartree's self consistent field method. Thus an electron with mass m in an ionic lattice may be described by an electronic wave function giving rise to an average charge distribution ρ. Treating this ρ as a static charge, it establishes a polarization in the lattice which can be calculated from the laws of electrostatics. This polarization gives rise to an attractive force on the electron, and to a potential energy of the order $-e^2/l\epsilon^*$ if the charge ρ extends over a spherical range whose radius is of the order l; ϵ^* is an effective dielectric constant. The restriction of the electron in space requires its de Broglie wave length to be of the order l so that its

kinetic energy is of the order $4\pi^2\hbar^2/2ml^2$. Minimizing the total energy E_0 with regard to l leads to

$$\frac{\partial}{\partial l}\left(-\frac{e^2}{l\epsilon^*}+\frac{4\pi^2\hbar^2}{2ml^2}\right)=0\,;\quad \frac{1}{l}=\frac{e^2m}{4\pi^2\hbar^2\epsilon^*}\,;\quad E_0\sim-\frac{1}{8\pi^2}\frac{e^4m}{\epsilon^{*2}\hbar^2}. \quad . \quad . \quad (1.1)$$

A procedure of this kind can be correct only if the time the electron requires to travel the distance l is short compared with the period of the vibrations of the lattice ; otherwise it is not permissible to treat ρ as a static charge distribution. This condition is not fulfilled in most cases of practical interest and it is necessary, therefore, to take account of the dynamic properties of the lattice. This is carried out in § 4. To obtain a qualitative picture it should be noticed that if $\omega/2\pi$ denotes the frequency of the lattice vibrations giving rise to polarization a point charge moving with velocity v polarizes the lattice to its full static value only at distances large compared with v/ω. Thus for sufficiently large distances the polarization potential decreases like a Coulomb field ; for distances less than v/ω it remains nearly constant. The interaction with the point charge is thus of the order $-e^2\omega/v\epsilon^*$. For an electron this holds only if its velocity is sufficiently large to make its de Broglie wave length h/mv shorter than v/ω. One may expect, therefore, that the lowest value the potential energy can have on the dynamic model is of the order $-e^2/d\epsilon^*$ where d is the length at which $h/mv=v/\omega$, i.e. $d=(h/m\omega)^{1/2}$. In § 4 it is shown, in fact, that the self energy on the dynamic model is of the order $-e^2/d\epsilon^*$.

The range of validity of the two methods will be discussed in § 6 ; it will also be shown then that the linking up of the two methods has not succeeded so far in view of certain difficulties inherent in the Hartree method.

Finally in § 7 questions of mobility will be discussed.

§ 2. Lattice Fields

The field of force acting on an electron brought into an insulating or semiconducting crystal can be derived from a periodic field of force. This latter will be defined as the field of force in an ideal crystal assuming that the atoms or ions of the lattice are undeformable and are rigidly fixed to the lattice points. F. Bloch (1928) has shown that an electron can move freely through such a periodic field of force though the connection between its energy and its momentum in general differs from that of a free electron in vacuum. In many cases, in particular for slow electrons in the conduction band, the electronic energy η_k is a quadratic function of the momentum $\hbar\mathbf{k}$ so that

$$\eta_k=\hbar^2k^2/2m, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.1)$$

as for free electrons although the effective mass m in general differs from the electronic mass m_{el}. This case (2.1) will always be assumed for the following.

Such an electron for which the periodic field of force has already been taken into account is subject to further forces resulting from deviations of the field of force from strict periodicity. These are due to displacements and deformations of atoms or ions (lattice defects will be disregarded here). Such displacements can be described by attaching a displacement vector to each lattice point. Moreover for electrons with sufficiently low energy the electronic wave function changes only very little in one lattice distance. Therefore the lattice can be treated as a continuum. The displacement vectors then become a continuous function of position, i.e. they form a vector field.

The present article is mainly concerned with isotropic ionic crystals. Here displacements leading to electric polarization have a much stronger interaction with electrons than other types of displacements. The latter will, therefore, be disregarded though their influence on the mobility will be mentioned in § 7. Let $\mathbf{P}(\mathbf{r})$ be the electric polarization at the position $\mathbf{r}$. Its interaction with the magnetic field of a slow electron is negligible compared with the electrostatic interaction. Therefore the rotational part of $\mathbf{P}$ is irrelevant so that the condition

$$\operatorname{curl} \mathbf{P}(\mathbf{r})=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.2)$$

can be imposed. Hence if electric displacement $\mathbf{D}(\mathbf{r})$ and electric field $\mathbf{E}(\mathbf{r})$ are introduced in the usual way by

$$\mathbf{D}(\mathbf{r})=\mathbf{E}(\mathbf{r})+4\pi\mathbf{P}(\mathbf{r}), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.3)$$

the only sources of $\mathbf{D}$ are free charges, i.e. in the case of a single free electron in the crystal,

$$\left.\begin{aligned} &\mathbf{D}(\mathbf{r}, \mathbf{r}_{el})=-\operatorname{grad}\frac{e}{|\mathbf{r}-\mathbf{r}_{el}|}, \\ \text{i.e.} \quad &\operatorname{curl}\mathbf{D}=0\,; \qquad \operatorname{div}\mathbf{D}=4\pi e\delta(\mathbf{r}-\mathbf{r}_{el}), \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (2.4)$$

where $\mathbf{r}_{el}$ is the electronic coordinate and $+e$ the electronic charge. In the absence of electrons $D=0$, therefore, so that $\mathbf{D}$ represents the 'external' field acting on the polarization of the crystal. Since $\mathbf{P}(\mathbf{r})\, d^3\mathbf{r}$ is the electric dipole moment of a volume $d^3\mathbf{r}$ around $\mathbf{r}$, clearly if a potential $\Phi(\mathbf{r})$ is defined by

$$4\pi\mathbf{P}(\mathbf{r})=\operatorname{grad}\Phi(\mathbf{r}), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.5)$$

the interaction energy between electron and lattice displacements is given by (using (2.4))

$$-\int D(\mathbf{r}, \mathbf{r}_{el})\mathbf{P}(\mathbf{r})\, d^3\mathbf{r}=\frac{1}{4\pi}\int \operatorname{grad}\frac{e}{|\mathbf{r}-\mathbf{r}_{el}|}\operatorname{grad}\Phi(\mathbf{r})\, d^3\mathbf{r}=e\Phi(\mathbf{r}_{el}). \quad . \quad (2.6)$$

Since the lattice polarization does not make any contribution to $\mathbf{D}$, it follows that $-4\pi\mathbf{P}$ represents the electric field due to the polarization, and $\Phi(\mathbf{r})$ is its potential.

It should be emphasized now that the polarization $\mathbf{P}$ is determined by the dynamic properties of the lattice. Thus while $\mathbf{D}$ is always given by (2.4) independent of the state of polarization of the lattice, the electric

field $\mathbf{E}$ depends on it, and can be found from (2.3) provided $\mathbf{P}$ has been determined from the dynamic equations. In reasonable approximation (cf. Fröhlich 1949, § 18) $\mathbf{P}$ can be considered as composed of two contributions

$$\mathbf{P}(\mathbf{r})=\mathbf{P}_0(\mathbf{r})+\mathbf{P}_{ir}(\mathbf{r}), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.7)$$

each of which satisfies a harmonic oscillator equation with frequencies $\omega_0/2\pi$ and $\omega/2\pi$ corresponding to optical absorption in the ultra violet ($\mathbf{P}_0$) and infra red ($\mathbf{P}_{ir}$) regions respectively. The former is mainly due to deformation and the latter due to displacement of ions. In the presence of an external field $\mathbf{D}$,

$$\ddot{\mathbf{P}}_{ir}(\mathbf{r})+\omega^2\mathbf{P}_{ir}(\mathbf{r})=\mathbf{D}(\mathbf{r}, \mathbf{r}_{el})/\gamma, \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.8)$$

$$\ddot{\mathbf{P}}_0(\mathbf{r})+\omega_0{}^2\mathbf{P}_0(\mathbf{r})=\mathbf{D}(\mathbf{r}, \mathbf{r}_{el})/\delta, \quad . \quad . \quad . \quad . \quad . \quad (2.9)$$

where γ and δ are constants closely connected with the energy due to the respective displacements (cf. eqn. (2.20)). The frequency $\omega\ 2\pi$ is correlated to the frequency $\omega_r/2\pi$ of residual rays by

$$\omega=(\epsilon/\epsilon_\infty)^{1/2}\omega_r, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.10)$$

where ϵ is the static dielectric constant and $\epsilon_\infty{}^{1/2}$ is the refractive index in the optical region before onset of ultra violet absorption.

To find γ and δ, consider the static case in which $\ddot{\mathbf{P}}=0$, i.e.

$$4\pi\mathbf{P}(\mathbf{r})=(\epsilon-1)\mathbf{E}(\mathbf{r})=(1-1/\epsilon)\mathbf{D}(\mathbf{r}), \quad . \quad . \quad . \quad . \quad (2.11)$$

making use of the static relation $\mathbf{D}=\epsilon\mathbf{E}$. Furthermore, the high frequency dielectric constant ϵ_∞ is defined by $\mathbf{D}=\epsilon_\infty\mathbf{E}$ on the assumption that the frequency $\omega_\infty/2\pi$ of the external field satisfies the condition $\omega_0 \gg \omega_\infty \gg \omega_{ir}$. In this case equations (2.8) and (2.9) show immediately that the infra red polarization can no longer follow and hence is negligible, $\mathbf{P}_{ir}\simeq 0$, but that the optical polarization $\mathbf{P}_0$ can follow nearly adiabatically so that $\ddot{\mathbf{P}}_0$ can be neglected as compared with $\omega_0{}^2\mathbf{P}_0$. Hence $\mathbf{P}_0$ has nearly the same value as in a static field of equal strength,

$$4\pi\mathbf{P}_0(\mathbf{r})=(1-1/\epsilon_\infty)\mathbf{D}(\mathbf{r}), \qquad \text{(static case)} \quad . \quad . \quad . \quad (2.12)$$

so that from (2.11), (2.12) and (2.7),

$$4\pi\mathbf{P}_{ir}(\mathbf{r})=\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right)\mathbf{D}(\mathbf{r}) \qquad \text{(static case)}. \quad . \quad . \quad . \quad (2.13)$$

Comparison of (2.8) and (2.9), for the static case ($\ddot{\mathbf{P}}_{ir}=0$, $\ddot{\mathbf{P}}_0=0$), with (2.12) and (2.13) respectively leads to

$$\frac{1}{\gamma}=\frac{\omega^2}{4\pi}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.14)$$

$$\frac{1}{\delta}=\frac{\omega_0{}^2}{4\pi}\left(1-\frac{1}{\epsilon_\infty}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.15)$$

The equations of motion (2.8) and (2.9) of the lattice field have to be supplemented by the equation of motion for the electron which follows

immediately with the help of the interaction energy (2.6). Thus classically

$$m\ddot{\mathbf{r}}_{el}=-e\,\mathrm{grad}\,\Phi(\mathbf{r}_{el}). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.16)$$

All three equations of motion can be derived from the Lagrangian

$$L=\tfrac{1}{2}\gamma\int[\dot{\mathbf{P}}_{ir}{}^2(\mathbf{r})-\omega^2\mathbf{P}_{ir}{}^2(\mathbf{r})]\,d^3\mathbf{r}+\tfrac{1}{2}\delta\int[\dot{\mathbf{P}}_0{}^2(\mathbf{r})-\omega_0{}^2\mathbf{P}_0{}^2(\mathbf{r})]\,d^3\mathbf{r}$$

$$+\int\mathbf{D}(\mathbf{r},\,\mathbf{r}_{el})[\mathbf{P}_0(\mathbf{r})+\mathbf{P}_{ir}(\mathbf{r})]\,d^3\mathbf{r}+\tfrac{1}{2}m\dot{\mathbf{r}}_{el}{}^2, \quad . \quad . \quad . \quad (2.17)$$

by putting

$$\delta L=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.18)$$

and treating the components of $\mathbf{r}_{el}$, $\mathbf{P}_0(\mathbf{r})$ and $\mathbf{P}_{ir}(\mathbf{r})$ as variables.† It will be noticed that in contrast to most other types of fields the spatial derivatives of the functions $\mathbf{P}_0(\mathbf{r})$ and $\mathbf{P}_{ir}(\mathbf{r})$ do not occur in the Lagrangian. Therefore in a straightforward way the components of these functions at different positions can be taken as different variables and (2.18) then yields immediately

$$\frac{d}{dt}\frac{\partial L}{\partial\dot{q}}-\frac{\partial L}{\partial q}=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.19)$$

if q represents any of these variables ; the differentiations with respect to them follow by considering the integrals as limits of sums. Equations (2.19) then lead to eqns. (2.16), and to (2.8) and (2.9) for all $\mathbf{r}$.

The moment conjugate to q is defined as $\partial L/\partial\dot{q}$ so that $\gamma\mathbf{P}_{ir}(\mathbf{r})$, $m\dot{\mathbf{r}}_{el}$ and $\delta\dot{\mathbf{P}}_0(\mathbf{r})$ are conjugate to $\mathbf{P}_{ir}(\mathbf{r})$, $\mathbf{r}_{el}$ and $\mathbf{P}_0(\mathbf{r})$ respectively. Hence the Hamiltonian is given by

$$H'=\sum_q\frac{\partial L}{\partial\dot{q}}\dot{q}-L=\tfrac{1}{2}\gamma\int[\dot{\mathbf{P}}_{ir}{}^2(\mathbf{r})+\omega^2\mathbf{P}_{ir}{}^2(\mathbf{r})]\,d^3\mathbf{r}$$

$$+\tfrac{1}{2}\delta\int[\dot{\mathbf{P}}_0{}^2(\mathbf{r})+\omega_0{}^2\mathbf{P}_0{}^2(\mathbf{r})]\,d^3\mathbf{r}$$

$$-\int D(\mathbf{r},\,\mathbf{r}_{el})[\mathbf{P}_{ir}(\mathbf{r})+\mathbf{P}_0(\mathbf{r})]\,d^3\mathbf{r}+\tfrac{1}{2}m\dot{\mathbf{r}}_{el}{}^2. \quad . \quad . \quad . \quad (2.20)$$

In the classical case of an electron at rest, when $\mathbf{P}_0$ and $\mathbf{P}_{ir}$ are given by (2.12) and (2.13), and $\mathbf{D}$ by (2.4), H' becomes infinite because the integrands in (2.20) diverge as $\mathbf{r}\to\mathbf{r}_{el}$. This arises from the disregard of the atomic structure of the substance. In a rough way this can be remedied by developing the polarization vectors in a Fourier series, and omitting terms with wave lengths shorter than a critical one whose wave length is of the order of the lattice distance.

In the quantum mechanical treatment this divergency does not occur as will be seen in §§ 4 and 5. Nevertheless, for simplicity, a classical argument will be used to eliminate the contributions of $\mathbf{P}_0$ to the

† In view of conditions (2.2), $\mathbf{P}_0$ and $\mathbf{P}_{ir}$ actually have only one available component each.

Hamiltonian. For this purpose it will be noticed that the force acting on a classical electron at rest must vanish, so that $\Phi(\mathbf{r})$ has a minimum at $\mathbf{r}=\mathbf{r}_{el}$ (provided the above mentioned 'cut off' at short waves is used) and $\mathbf{P}(\mathbf{r}_{el})$ vanishes. For a moving electron, however, $\mathbf{P}(\mathbf{r}_{el})$ does not vanish because of the time dependent terms in the equations of motion for $\mathbf{P}_{ir}(\mathbf{r})$ and $\mathbf{P}_0(\mathbf{r})$. Inspection of eqns. (2.8) and (2.9) for a time dependent $\mathbf{D}$ shows that for slowly moving electrons $\mathbf{P}_0(\mathbf{r}_{el})$ remains nearly zero and at any rate is much smaller than $\mathbf{P}_{ir}(\mathbf{r}_{el})$ because $\omega_0{}^2$ is very large compared with $\omega_{ir}{}^2$. Physically it means that the deformation of the ions giving rise to $\mathbf{P}_0$ has a much smaller inertia than the displacement of ions (leading to $\mathbf{P}_{ir}$). It follows from the above that all terms in H' containing $\mathbf{P}_0(\mathbf{r})$ depend only very insensitively on the velocity of the electron. For small velocity the quantity

$$C=\int\{\tfrac{1}{2}\delta[\dot{\mathbf{P}}_0{}^2(\mathbf{r})+\omega_0{}^2\mathbf{P}_0{}^2(\mathbf{r})]-D(\mathbf{r},\,\mathbf{r}_{el})\mathbf{P}_0(\mathbf{r})\}\,d^3\mathbf{r}, \quad . \quad . \quad (2.21)$$

can then be considered as a constant and the Hamiltonian of a slow electron in an ionic lattice field can then be written as

$$H=H'-C=\int\{\tfrac{1}{2}\gamma[\dot{\mathbf{P}}_{ir}{}^2(\mathbf{r})+\omega^2\mathbf{P}_{ir}{}^2(\mathbf{r})]-\mathbf{D}(\mathbf{r},\,\mathbf{r}_{el})\mathbf{P}_{ir}(\mathbf{r})\}\,d^3\mathbf{r}+\tfrac{1}{2}m\dot{\mathbf{r}}_{el}{}^2. \quad . \quad . \quad . \quad (2.22)$$

It is useful to introduce here a complex field vector $\mathbf{B}(\mathbf{r})$ such that

$$\mathbf{B}(\mathbf{r})=\left(\frac{\gamma\omega}{2\hbar}\right)^{1/2}\left(\mathbf{P}_{ir}(\mathbf{r})+\frac{i}{\omega}\dot{\mathbf{P}}_{ir}(\mathbf{r})\right), \quad . \quad . \quad . \quad . \quad . \quad (2.23)$$

$$\mathbf{B}^+(\mathbf{r})=\left(\frac{\gamma\omega}{2\hbar}\right)^{1/2}\left(\mathbf{P}_{ir}(\mathbf{r})-\frac{i}{\omega}\dot{\mathbf{P}}_{ir}(\mathbf{r})\right), \quad . \quad . \quad . \quad (2.23^+)$$

$$\operatorname{curl}\mathbf{B}=\operatorname{curl}\mathbf{B}^+=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.24)$$

or

$$\mathbf{P}_{ir}(\mathbf{r})=\left(\frac{\hbar}{2\gamma\omega}\right)^{1/2}[\mathbf{B}^+(\mathbf{r})+\mathbf{B}(\mathbf{r})], \quad . \quad . \quad . \quad . \quad . \quad (2.25)$$

$$\dot{\mathbf{P}}_{ir}(\mathbf{r})=\left(\frac{\hbar\omega}{2\gamma}\right)^{1/2}i[\mathbf{B}^+(\mathbf{r})-\mathbf{B}(\mathbf{r})]. \quad . \quad . \quad . \quad . \quad . \quad (2.26)$$

Then with (2.22),

$$H=\hbar\omega\int\mathbf{B}^+(\mathbf{r})\mathbf{B}(\mathbf{r})\,d^3\mathbf{r}-\left(\frac{\hbar}{2\gamma\omega}\right)^{1/2}\int\mathbf{D}(\mathbf{r},\,\mathbf{r}_{el})[\mathbf{B}^+(\mathbf{r})+\mathbf{B}(\mathbf{r})]\,d^3\mathbf{r}+\frac{\mathbf{p}_{el}{}^2}{2m}, \quad . \quad . \quad . \quad (2.27)$$

where $\mathbf{p}_{el}=m\dot{\mathbf{r}}_{el}$. The constants in the definition of $\mathbf{B}$ have been chosen in view of the following quantization. For this purpose $\mathbf{B}(\mathbf{r})$ will be subjected to a periodic boundary condition over a cube of volume $V=L^3$, i.e. considering (2.24),

$$\mathbf{B}(\mathbf{r})=\sum_{\mathbf{w}}\frac{\mathbf{w}}{w}b_{\mathbf{w}}\frac{\exp(i\mathbf{wr})}{V^{1/2}}, \quad \mathbf{B}^+(\mathbf{r})=\sum_{\mathbf{w}}\frac{\mathbf{w}}{w}b_{\mathbf{w}}{}^+\frac{\exp(-i\mathbf{wr})}{V^{1/2}}, \quad (2.28)$$

where the coefficients w_i of the wave vector $\mathbf{w}$ satisfy

$$w_i = \frac{2\pi}{L} n_i, \quad n_i = 0, \pm 1, \pm 2, \ldots \quad (2.29)$$

Also defining (cf. (2.5)) $\Phi_{ir}(\mathbf{r})$ by

$$4\pi \mathbf{P}_{ir}(\mathbf{r}) = \text{grad}\, \Phi_{ir}(\mathbf{r}), \quad (2.30)$$

it follows that (cf. (2.6)) the interaction energy is given by

$$-\int \mathbf{D}(\mathbf{r}, \mathbf{r}_{el}) \mathbf{P}_{ir}(\mathbf{r})\, d^3\mathbf{r} = e\Phi_{ir}(\mathbf{r}_{el}), \quad (2.31)$$

or using (2.30), (2.25) and (2.28), by

$$e\Phi_{ir}(\mathbf{r}_{el}) = 4\pi i \left(\frac{e^2\hbar}{2\gamma\omega V}\right)^{1/2} \sum_{\mathbf{w}} \frac{1}{w} [b_{\mathbf{w}}^{+} \exp(-i\mathbf{w}\mathbf{r}_{el}) - b_{\mathbf{w}} \exp(i\mathbf{w}\mathbf{r}_{el})]. \quad (2.32)$$

Hence with (2.27),

$$H = \frac{p_{el}^2}{2m} + \hbar\omega \sum_{\mathbf{w}} b_{\mathbf{w}}^{+} b_{\mathbf{w}} + 4\pi i \left(\frac{e^2\hbar}{2\gamma\omega V}\right)^{1/2} \sum_{\mathbf{w}} \frac{1}{w} [b_{\mathbf{w}}^{+} \exp(-i\mathbf{w}\mathbf{r}_{el}) - b_{\mathbf{w}} \exp(i\mathbf{w}\mathbf{r}_{el})]. \quad (2.33)$$

It will be noticed that (2.32) holds only if $w \neq 0$, and in fact terms with $w = 0$ must be excluded from all the sums. To justify this it should be realized that strictly speaking the case of a single electron in an otherwise neutral lattice cannot be subjected to a periodic boundary condition. For the volume V contains an electron, i.e. the charge e, a situation which cannot be continued periodically in view of the long range of repulsive Coulomb interaction. To avoid this difficulty an opposite charge, $-e$, should be distributed homogeneously over the volume V. This just compensates the component $w = 0$ of the electronic charge, and hence there is no $w = 0$ term in the interaction energy.

For quantization it will be remembered that the components of $\mathbf{P}_{ir}(\mathbf{r})$ and $\gamma \dot{\mathbf{P}}_{ir}(\mathbf{r})$ are conjugate variables. Making use of (2.25), (2.26), (2.28) and (2.2), this means that $(\hbar/2\gamma\omega)^{1/2}(b_{\mathbf{w}}^{+} + b_{\mathbf{w}})$ and $\gamma(\hbar\omega/2\gamma)^{1/2} i(b_{\mathbf{w}}^{+} - b_{\mathbf{w}})$ are conjugate. Thus using the notation $[u, v] = uv - vu$, quantization according to Bose statistics demands that

$$\tfrac{1}{2} i\hbar [b_{\mathbf{w}}^{+} + b_{\mathbf{w}}, b_{\mathbf{w}}^{+} - b_{\mathbf{w}}] = i\hbar, \quad (2.34)$$

and that commutators involving different $\mathbf{w}$'s vanish. It then follows that

$$[b_{\mathbf{w}}, b_{\mathbf{w}'}^{+}] = \delta_{\mathbf{w},\mathbf{w}'}; \qquad [b_{\mathbf{w}}, b_{\mathbf{w}'}] = 0. \quad (2.35)$$

Also, in the Hamiltonian, the terms with $b_{\mathbf{w}}^{+} b_{\mathbf{w}}$ should by symmetrized and hence be replaced by $\frac{1}{2}(b_{\mathbf{w}}^{+} b_{\mathbf{w}} + b_{\mathbf{w}} b_{\mathbf{w}}^{+}) = b_{\mathbf{w}}^{+} b_{\mathbf{w}} + \frac{1}{2}$, using (2.35). This means that (2.33) should be augmented by $\hbar\omega \Sigma_{\mathbf{w}} \frac{1}{2}$ which is a constant, representing the zero point energy of the oscillating ions. This term can be omitted, however, because it would simply change the zero of the total

energy. Also, in the usual way, the components p_{jel} of $\mathbf{p}_{el}$ must be replaced by the operator $(\hbar/i)\partial/\partial r_{jel}$, i.e.

$$[p_{jel}, r_{kel}] = -i\hbar\delta_{jk}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.36)$$

It is useful at this stage to introduce dimensionless variables. because this shows that the problem presented by the Hamiltonian H. eqn. (2.33), essentially depends on a single parameter, the interaction parameter which is defined by

$$\alpha = \frac{2\pi e^2 u}{\hbar\gamma\omega^3} = \frac{1}{2}\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon}\right)\frac{e^2 u}{\hbar\omega}. \quad . \quad . \quad . \quad . \quad . \quad (2.37)$$

Use has been made of (2.14) ; u is an inverse length defined by

$$\frac{\hbar^2 u^2}{2m} = \hbar\omega, \qquad \text{i.e. } u = \left(\frac{2m\omega}{\hbar}\right)^{1/2} . \quad . \quad . \quad . \quad . \quad (2.38)$$

Define now the dimensionless quantities

$$\mathbf{x} = u\mathbf{r}_{el}, \qquad \mathbf{v} = \mathbf{w}/u, \qquad S = Vu^3, \quad . \quad . \quad . \quad . \quad (2.39)$$

and express the Hamiltonian in terms of $\hbar\omega$. Then from (2.33), using (2.37)–(2.39),

$$F = \frac{H}{\hbar\omega}$$
$$= -\sum_1^3 \frac{\partial^2}{\partial x_i^2} + \sum_{\mathbf{v}} b_{\mathbf{v}}^+ b_{\mathbf{v}} + i\left(\frac{4\pi\alpha}{S}\right)^{1/2} \sum_{\mathbf{v}} \frac{1}{v}[b_{\mathbf{v}}^+ \exp(-i\mathbf{vx}) - b_{\mathbf{v}} \exp(i\mathbf{vx})]. \quad (2.40)$$

where x_i are the components of $\mathbf{x}$.

§ 3. General Properties

The aim of the present section is to discuss some general properties of the polarization field $\mathbf{P}_{ir}(\mathbf{r})$ or $\mathbf{B}(\mathbf{r})$. Consider first the field in the absence of an electron. Its Hamiltonian H_{field} follows from (2.33) by omitting the electronic and the interaction terms, i.e.

$$H_{\text{field}} = \Sigma H_{\mathbf{w}} ; \quad H_{\mathbf{w}} = \hbar\omega b_{\mathbf{w}}^+ b_{\mathbf{w}}. \quad . \quad . \quad . \quad . \quad . \quad (3.1)$$

Here $H_{\mathbf{w}}$ will be recognized as the Hamiltonian of a harmonic oscillator with frequency $\omega/2\pi$, coordinate $q_{\mathbf{w}}$ and momentum $p_{\mathbf{w}}$ where

$$q_{\mathbf{w}} = \left(\frac{\hbar}{2M\omega}\right)^{1/2} (b_{\mathbf{w}}^+ + b_{\mathbf{w}}) ; \quad p_{\mathbf{w}} = i\left(\frac{M\hbar\omega}{2}\right)^{1/2} (b_{\mathbf{w}}^+ - b_{\mathbf{w}}) \quad . \quad (3.2)$$

and M is its mass. For then

$$H_{\mathbf{w}} = \frac{p_{\mathbf{w}}^2}{2M} + \tfrac{1}{2}M\omega^2 q_{\mathbf{w}}^2 - \tfrac{1}{2}\hbar\omega. \quad . \quad . \quad . \quad . \quad . \quad (3.3)$$

From the theory of the harmonic oscillator it follows that the eigenvalues of $b_{\mathbf{w}}^+ b_{\mathbf{w}}$ are positive integers, or zero, i.e.

$$b_{\mathbf{w}}^+ b_{\mathbf{w}} \,|\, n_{\mathbf{w}}\rangle = n \,|\, n_{\mathbf{w}}\rangle, \qquad n = 0, 1, 2, \ldots, \quad . \quad . \quad . \quad (3.4)$$

where $|n_{\mathbf{w}}\rangle$ denotes a normalized eigenfunction. Also it is easily proved that (n is a positive integer)

$$b_{\mathbf{w}}{}^{+n}|0_{\mathbf{w}}\rangle=(n!)^{1/2}|n_{\mathbf{w}}\rangle, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.5)$$

and

$$b_{\mathbf{w}}{}^{n}|n_{\mathbf{w}}\rangle=(n!)^{1/2}|0_{\mathbf{w}}\rangle\,; \quad b_{\mathbf{w}}|0_{\mathbf{w}}\rangle=0. \quad . \quad . \quad . \quad . \quad (3.6)$$

In view of these properties the operators $b_{\mathbf{w}}{}^{+}$ and $b_{\mathbf{w}}$ are denoted as creation and annihilation operators, respectively.

The field is thus represented by a set of oscillators $\mathbf{w}$, all having the same frequency $\omega/2\pi$. Alternatively the field can be considered as represented by polarization quanta of different wave numbers $\mathbf{w}$, but all having the same energy $\hbar\omega$, independent of $\mathbf{w}$. The corresponding classical polarization waves have thus equal frequency, independent of the wave number, a property which follows also directly from eqn. (2.8). Hence

$$\partial\omega/\partial\mathbf{w}=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.7)$$

which means that the group velocity of polarization waves vanishes.

One would conclude from this that the field cannot transport momentum. In field theory the quantity $\hbar\mathbf{w}$ is, however, usually denoted as momentum of a quantum because in the presence of an electron the operator

$$\hbar\mathbf{W}=\sum_{\mathbf{w}}\hbar\mathbf{w}b_{\mathbf{w}}{}^{+}b_{\mathbf{w}}+\mathbf{p}_{el} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.8)$$

is a constant of motion, i.e.

$$[\mathbf{W}, H]=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.9)$$

Equation (3.9) follows immediately from (3.8) and (2.33) using the commutators (2.35) and (2.36). For clearly $p_{el}{}^{2}$ and $b_{\mathbf{w}}{}^{+}b_{\mathbf{w}}$ commute both with $\mathbf{W}$ whereas

$$[b_{\mathbf{w}}{}^{+}b_{\mathbf{w}}, b_{\mathbf{w}}{}^{+}]=b_{\mathbf{w}}{}^{+}, \qquad [p_{el}, \exp(-i\mathbf{w}\mathbf{r}_{el})]=-\hbar\mathbf{w}\exp(-i\mathbf{w}\mathbf{r}_{el}), \text{ etc.}$$
$$. \quad . \quad . \quad . \quad (3.10)$$

Hence the commutator of $\mathbf{p}_{el}$ with the interaction term in the Hamiltonian just cancels that of $\Sigma\hbar\mathbf{w}b_{\mathbf{w}}{}^{+}b_{\mathbf{w}}$. In spite of (3.9) it will be shown presently that a polarization quantum does not carry any momentum. In fact through discussions with B. Szigeti and S. Zienau it was realized that in any lattice field the quanta cannot carry momentum because the individual atoms or ions of the lattice oscillate around their respective equilibrium positions, and hence on an average have zero momentum. The momentum of the lattice is always carried by its centre of gravity as will be demonstrated presently. Hence (3.9) does not represent the momentum law but the law of conservation of the total wave vector.

To show this it will be necessary to introduce the coordinate $\mathbf{r}_c$ of the centre of gravity of the lattice and treat it as a dynamic variable. This means that the developments of § 2, in particular the equations of motion, the Lagrangian and the Hamiltonian require modifications because in § 2 it was tacitly assumed that the centre of gravity is at rest. To carry

this through let $\mathbf{r}'$ be the positional vector in a fixed coordinate system relative to which the centre of gravity has the coordinate $\mathbf{r}_c$, and the electron has the coordinate $\mathbf{r}_{el}'$, i.e.

$$\mathbf{r}'=\mathbf{r}+\mathbf{r}_c, \qquad \mathbf{r}_{el}'=\mathbf{r}_{el}+\mathbf{r}_c. \quad . \quad . \quad . \quad . \quad . \quad (3.11)$$

Hence the polarization vector $\mathbf{P}_{ir}(\mathbf{r}, t)$ has to be replaced by

$$\mathbf{P}_{ir}(\mathbf{r}'-\mathbf{r}_c, t)=\mathbf{P}'(\mathbf{r}', t),$$

and

$$\frac{\partial \mathbf{P}_{ir}(\mathbf{r}, t)}{\partial t} \quad \text{by} \quad \frac{\partial \mathbf{P}'(\mathbf{r}', t)}{\partial t}+\left(\dot{\mathbf{r}}_c \frac{\partial}{\partial \mathbf{r}'}\right)\mathbf{P}'(\mathbf{r}', t). \quad . \quad . \quad (3.12)$$

With these replacements, the new Lagrangian L' is obtained from the field terms of (2.17) (the $\mathbf{P}_0$-terms will be omitted) by adding the kinetic energies of electron and centre of gravity. Thus introducing the operator

$$\partial=\partial(t, \mathbf{r}', \dot{\mathbf{r}}_c)=\frac{\partial}{\partial t}+\left(\dot{\mathbf{r}}_c \frac{\partial}{\partial \mathbf{r}'}\right), \quad . \quad . \quad . \quad . \quad . \quad (3.13)$$

one finds (NM is the mass of the lattice)

$$L'=\int \{\tfrac{1}{2}\gamma(\partial \mathbf{P}')^2-\tfrac{1}{2}\gamma\omega^2\mathbf{P}'^2+\mathbf{D}\mathbf{P}'\}\, d^3\mathbf{r}'+\tfrac{1}{2}m\dot{\mathbf{r}}_{el}'^2+\tfrac{1}{2}NM\dot{\mathbf{r}}_c^2. \quad . \quad (3.14)$$

It will be noted that $\mathbf{D}(\mathbf{r}, \mathbf{r}_{el})=\mathbf{D}(\mathbf{r}', \mathbf{r}_{el}')$ because this function depends on $|\mathbf{r}-\mathbf{r}_{el}|$ only. The variables in L' are the components of $\mathbf{r}_c$, $\mathbf{r}_{el}'$ and $\mathbf{P}'(\mathbf{r}', t)$ the latter for all value of $\mathbf{r}'$. The conjugate momenta are given by

$$p_{ci}=\frac{\partial L'}{\partial \dot{\mathbf{r}}_{ci}}=NM\dot{\mathbf{r}}_{ci}+\gamma\int \partial \mathbf{P}' \frac{\partial}{\partial r_i'}\mathbf{P}'\, d^3\mathbf{r}', \quad . \quad . \quad . \quad (3.15)$$

$$\mathbf{p}_{el}'=\frac{\partial L'}{\partial \dot{\mathbf{r}}_{el}'}=m\dot{\mathbf{r}}_{el}' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.16)$$

and

$$\mathbf{Q}'=\frac{\partial L'}{\partial \dot{\mathbf{P}}'}=\gamma\partial(t, \mathbf{r}', \dot{\mathbf{r}}_c)\mathbf{P}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.17)$$

The equations of motion follow with the help of (2.19). Since spatial derivatives of $\mathbf{P}'$ now occur under the integral, the Lagrangian derivative with regard to $\mathbf{P}'$ has to be taken, i.e.

$$\left(\frac{\partial}{\partial q}\right)_{q=P_j'} \to \frac{\partial}{\partial P_j'}-\sum_i \frac{\partial}{\partial r_i'}\frac{\partial}{\partial(\partial P_j'/\partial r_i')}. \quad . \quad . \quad . \quad . \quad (3.18)$$

Hence one finds for $q \to \mathbf{r}_{el}'$, with (2.4),

$$m\ddot{\mathbf{r}}_{el}'=\int \frac{\partial \mathbf{D}(|\mathbf{r}'-\mathbf{r}_{el}'|)}{\partial \mathbf{r}_{el}'}\mathbf{P}'(\mathbf{r}')\, d^3\mathbf{r}'=-\int \frac{\partial \mathbf{D}'}{\partial \mathbf{r}'}\mathbf{P}'\, d^3\mathbf{r}'$$

$$=-4\pi e\mathbf{P}'(\mathbf{r}_{el}')=-e \operatorname{grad} \Phi'(\mathbf{r}_{el}'), \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.19)$$

where

$$e\Phi'(\mathbf{r}_{el}')=-\int D(|\mathbf{r}'-\mathbf{r}_{el}'|)\mathbf{P}'(\mathbf{r}')\, d^3\mathbf{r}', \quad . \quad . \quad . \quad . \quad (3.20)$$

denotes the interaction energy between field and electron.

For $q \to r_{ci}$, (2.19) yields

$$NM\ddot{r}_{ci}+\gamma \frac{\partial}{\partial t}\int \partial(t, \mathbf{r}', \dot{\mathbf{r}}_c)\mathbf{P}'(\mathbf{r}') \frac{\partial}{\partial r_i'} \mathbf{P}'(\mathbf{r}')\, d^3\mathbf{r}'=0, \quad . \quad . \quad (3.21)$$

and for $q \to \mathbf{P}'$,

$$\partial^2(t, \mathbf{r}', \dot{\mathbf{r}}_c)\mathbf{P}'(\mathbf{r}')+\omega^2\mathbf{P}'(\mathbf{r}')=\mathbf{D}(\mathbf{r}', \mathbf{r}_{el}')/\gamma. \quad . \quad . \quad . \quad (3.22)$$

The latter follows also directly from (2.8) with the replacement (3.13). Clearly if $\dot{\mathbf{r}}_{el}'=\dot{\mathbf{r}}_c$, then $\mathbf{P}'=\mathbf{D}(|\mathbf{r}'-\mathbf{r}_{el}'|)/\gamma\omega^2$ is a particular solution, as required. The solutions of the homogeneous equation (3.22) (i.e. of (3.22) with $\mathbf{D}=0$), can be written as a superposition of plane waves such that for one of them

$$\mathbf{P}_{\mathbf{w}}' \propto \exp(i\mathbf{w}\mathbf{r}'-i\omega' t), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.23)$$

where

$$\omega'=\omega+\dot{\mathbf{r}}_i\mathbf{w} \qquad \text{or} \qquad \partial\omega'/\partial\mathbf{w}=\dot{\mathbf{r}}_c, \quad . \quad . \quad . \quad (3.24)$$

the latter with the use of (3.7). Thus the group velocity is now equal to the velocity of the centre of gravity, as required.

To discuss (3.21) it will be noticed that the integral occurring in this equation can be transformed. For with the use of (3.13),

$$\frac{\partial}{\partial t}\int \partial(t, \mathbf{r}', \dot{\mathbf{r}}_c)\mathbf{P}'(\mathbf{r}') \frac{\partial}{\partial r_i'} \mathbf{P}'(\mathbf{r}')\, d^3\mathbf{r}$$

$$= \int\left[\left\{\partial^2\mathbf{P}'-\left(\dot{\mathbf{r}}_c\frac{\partial}{\partial \mathbf{r}'}\right)\left(\dot{\mathbf{P}}'+\left(\dot{\mathbf{r}}_c\frac{\partial}{\partial \mathbf{r}'}\right)\mathbf{P}'\right)\right\}\frac{\partial \mathbf{P}'}{\partial r_i'}+\left(\dot{\mathbf{P}}'+\left(\dot{\mathbf{r}}_c\frac{\partial}{\partial \mathbf{r}'}\right)\mathbf{P}'\right)\frac{\partial \dot{\mathbf{P}}}{\partial r_i'}\right]d^3\mathbf{r}'.$$

. . . . (3.25)

Now through integration in parts it is found that

$$\int \dot{\mathbf{P}}'\frac{\partial \dot{\mathbf{P}}'}{\partial r_i'}d^3\mathbf{r}'=-\int \dot{\mathbf{P}}'\frac{\partial \dot{P}'}{\partial r_i'}d^3\mathbf{r}'=0\,; \quad \text{similarly} \int \mathbf{P}'\frac{\partial \dot{\mathbf{P}}'}{\partial r_i'}d^3\mathbf{r}'=0$$

and

$$\int \frac{\partial^2\mathbf{P}'}{\partial r_k'\partial r_j'}\frac{\partial \mathbf{P}'}{\partial r_i'}d^3\mathbf{r}'=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.26)$$

because all surface integrals vanish. For this reason, also

$$\int\left(\dot{\mathbf{r}}_c\frac{\partial}{\partial \mathbf{r}'}\right)\left(\mathbf{P}'\frac{\partial \dot{\mathbf{P}}'}{\partial r_i'}-\dot{\mathbf{P}}'\frac{\partial \mathbf{P}'}{\partial r_i'}\right)d^3\mathbf{r}'=0. \quad . \quad . \quad . \quad . \quad (3.27)$$

Thus on the right-hand side of (3.25) all terms vanish except $\partial^2\mathbf{P}'$. Making use of (3.22) and again of (3.26), and of curl $\mathbf{P}'=0$,

$$\int \partial^2\mathbf{P}'\frac{\partial \mathbf{P}'}{\partial r_i'}d^3\mathbf{r}'=-\omega^2\int \mathbf{P}'\frac{\partial \mathbf{P}'}{\partial r_i'}d^3\mathbf{r}'+\frac{1}{\gamma}\int \mathbf{D}\frac{\partial \mathbf{P}'}{\partial r_i'}d^3\mathbf{r}'=\frac{1}{\gamma}\int\left(\mathbf{D}\frac{\partial}{\partial \mathbf{r}'}\right)P_i'd^3\mathbf{r}'$$

$$=-\frac{1}{\gamma}\int\frac{\partial \mathbf{D}}{\partial \mathbf{r}'}\mathbf{P}_i'=-4\pi e P_i'(\mathbf{r}_{el}'), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.28)$$

the latter with the use of (2.4). Thus from (3.21) and the above,

$$NM\ddot{\mathbf{r}}_c=4\pi e\mathbf{P}'(\mathbf{r}_{el}'). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.29)$$

Comparison with (3.19) shows that

$$\frac{d\mathbf{p}}{dt}=0, \quad \text{where} \quad \mathbf{p}=m\dot{\mathbf{r}}_{el}+NM\dot{\mathbf{r}}_c, \quad . \quad . \quad . \quad . \quad (3.30)$$

which represents the momentum law on the usual definition of centre of gravity momentum by $NM\dot{\mathbf{r}}_c$. This definition differs from that of $\mathbf{p}_c$ by (3.15) ; this is a pure formality required by the Lagrangian formalism in view of the occurrence of $\dot{\mathbf{r}}_c$ in the field terms. In eqn. (3.30) all properties of polarization waves have been eliminated. This equation thus shows that any change of the electronic momentum by interaction with polarization waves is compensated by an opposite change in $NM\dot{\mathbf{r}}_c$. i.e. by the centre of gravity momentum (in the conventional definition) of the lattice.

The new Hamiltonian now follows in the usual way from (3.14) using (3.15)–(3.17),

$$H'=\sum_q \dot{q}\frac{\partial L'}{\partial \dot{q}}-L'=\tfrac{1}{2}NM\dot{\mathbf{r}}_c{}^2+\tfrac{1}{2}m\dot{\mathbf{r}}_{el}{}'^2+\int\{\tfrac{1}{2}\gamma[(\partial\mathbf{P}')^2+\omega^2\mathbf{P}'^2]-\mathbf{D}\mathbf{P}'\}d^3\mathbf{r}'. \quad . \quad . \quad . \quad (3.31)$$

Again a complex quantity $\mathbf{B}'(\mathbf{r}')$ can be introduced similar to (2.23)–(2.26) and (2.28) by

$$\mathbf{B}'(\mathbf{r}')=\left(\frac{\gamma\omega}{2\hbar}\right)^{1/2}\left(\mathbf{P}'+\frac{i}{\omega}\partial\mathbf{P}'\right)=\sum_{\mathbf{w}}\frac{\mathbf{w}}{w}b_{\mathbf{w}}{}'\frac{\exp(i\mathbf{w}\mathbf{r}')}{V^{1/2}}, \quad . \quad (3.32)$$

or

$$\mathbf{P}'(\mathbf{r}')=\left(\frac{\hbar}{2\gamma\omega}\right)^{1/2}[\mathbf{B}'^{+}(\mathbf{r}')+\mathbf{B}'(\mathbf{r}')], \quad . \quad . \quad . \quad . \quad . \quad (3.33)$$

$$\partial\mathbf{P}'(\mathbf{r}')=\left(\frac{\hbar\omega}{2\gamma}\right)^{1/2}i[\mathbf{B}'^{+}(\mathbf{r}')-\mathbf{B}'(\mathbf{r}')]. \quad . \quad . \quad . \quad . \quad . \quad (3.34)$$

Introducing this into (3.15), the momentum $\mathbf{p}_c$ is found to be

$$\mathbf{p}_c=NM\mathbf{r}_c-\sum\hbar\mathbf{w}b_{\mathbf{w}}{}'^{+}b_{\mathbf{w}}{}'. \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.35)$$

Hence from (3.31) using (3.35), (3.16) and (3.32)–(3.34) by a procedure similar to that leading to (2.33),

$$H'=\frac{1}{2NM}(\mathbf{p}_c+\sum_{\mathbf{w}}\hbar\mathbf{w}b_{\mathbf{w}}{}'^{+}b_{\mathbf{w}}{}')^2+\frac{\mathbf{p}_{el}{}'^2}{2m}+\hbar\omega\sum_{\mathbf{w}}b_{\mathbf{w}}{}'^{+}b_{\mathbf{w}}{}'$$
$$+4\pi i\left(\frac{e^2\hbar}{2\gamma\omega V}\right)^{1/2}\sum_{\mathbf{w}}\frac{1}{w}[b_{\mathbf{w}}{}'^{+}\exp(-i\mathbf{w}\mathbf{r}_{el}{}')-b_{\mathbf{w}}{}'\exp(i\mathbf{w}\mathbf{r}_{el}{}')]. \quad . \quad (3.36)$$

Quantization can now be carried out in the usual way by introduction of the commutators

$$[b_{\mathbf{w}}{}', b_{\mathbf{w}}{}'^{+}]=\delta_{\mathbf{w},\mathbf{w}'}, \qquad [p'_{elj}, r'_{elj}]=-i\hbar, \qquad [p_{cj}, r_{cj}]=-i\hbar. \quad . \quad (3.37)$$

Thus the conservation law for the total wave number holds as in (3.8) and (3.9), i.e.

$$[\mathbf{W}', H']=0 \qquad \text{with} \qquad \hbar\mathbf{W}'=\mathbf{p}_{el}{}'+\sum_{\mathbf{w}}\hbar\mathbf{w}b_{\mathbf{w}}{}'^{+}b_{\mathbf{w}}{}'. \quad . \quad (3.38)$$

In addition, however, $\mathbf{p}_c$ is also conserved because

$$[\mathbf{p}_c, H']=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.39)$$

This follows because H' is independent of $\mathbf{r}_c$. It must be noted that

$$\dot{\mathbf{r}}_c=\frac{\hbar}{i}[H', \mathbf{r}_c]=\frac{1}{NM}(\mathbf{p}_c+\sum_{\mathbf{w}}\hbar\mathbf{w}b_{\mathbf{w}}{}'^{+}b_{\mathbf{w}}{}'), \quad . \quad . \quad . \quad (3.40)$$

so that with the use of (3.38), eqn. (3.39) is found to be equivalent with

$$[\mathbf{p}, H']=0 \qquad \text{where} \qquad \mathbf{p}=\mathbf{p}_{el}+NM\dot{\mathbf{r}}_c \quad . \quad . \quad . \quad (3.41)$$

a conservation law already found previously (3.30) for the classical case.

On the usual definition, $\mathbf{p}$ represents the total momentum of the system. Its conservation is clearly a law different from the conservation of the total wave number $\mathbf{W}'$.

In all practical cases, NM is so large that the first term in the Hamiltonian can be neglected, i.e. it is sufficient to deal with (2.33) or (2.40). The main purpose of the above discussion was to clarify the physical meaning of the conservation laws, and to show how momentum is conserved even though polarization quanta do not carry momentum.

§ 4. Stationary Solutions I

The present section deals mainly with the derivation of approximate wave functions and eigenvalues of the Hamiltonian (2.33) or (2.40). The problem is one of finding the self energy of a single electron in the polarization field. The simplest method to find solutions is application of perturbation theory (cf. Fröhlich, Pelzer and Zienau 1950). This method holds only for weak interaction, i.e. when $\alpha<1$.

4.1. *Perturbation Theory*

Consider first the solutions of the Hamiltonian

$$F_0=-\sum_1^3 \frac{\partial^2}{\partial x_i^2}+\sum_{\nu} b_{\nu}^{+}b_{\nu} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.1)$$

obtained from (2.40) by omitting the interaction terms proportional to the interaction parameter α. The wave functions are products of a free electron wave function, and oscillator wave functions of the type defined in (3.4),

$$|\,\mathbf{k}\,;\, n_{\nu_1}, n_{\nu_2}, \ldots n_{\nu} \ldots\rangle = S^{-1/2} \exp(i\mathbf{qx}) \prod_{\nu} |\,n_{\nu}\rangle, \quad . \quad . \quad (4.2)$$

where (cf. (2.38))

$$\mathbf{k}=\mathbf{q}+\sum_{\nu}\boldsymbol{\nu} n_{\nu}\,; \quad u\mathbf{k}=\mathbf{W} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.3)$$

is the total wave vector in dimensionless form (cf. (2.39)), and

$$b_{\nu}^{+}b_{\nu}|\,n_{\nu}\rangle = n_{\nu}|\,n_{\nu}\rangle \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.4)$$

similar to (3.4). Clearly $\mathbf{q}$ is determined by (4.3) in terms of $\mathbf{k}$ and the n_{ν}'s and need thus not be explicitly included in the parameters defining the wave function. The corresponding energy ϵ_0 (multiples of $\hbar\omega$) satisfying

$$F_0|\,\mathbf{k}\,;\, \ldots n_{\nu} \ldots\rangle = \epsilon_0(\mathbf{k}\,;\, \ldots n_{\nu} \ldots)|\,\mathbf{k}\,;\, \ldots n_{\nu} \ldots\rangle \quad . \quad (4.5)$$

is given by

$$\epsilon_0(\mathbf{k}\,;\, \ldots n_{\nu} \ldots)=q^2+\sum_{\nu} n_{\nu}=(\mathbf{k}-\sum_{\nu}\boldsymbol{\nu} n_{\nu})^2+\sum_{\nu} n_{\nu}. \quad . \quad (4.6)$$

The interaction term (cf. (2.40))

$$F_{\text{int}}=F-F_0=i\left(\frac{4\pi\alpha}{S}\right)^{1/2}\sum_{\mathbf{v}}\frac{1}{v}[b_{\mathbf{v}}^{+}\exp(-i\mathbf{vx})-b_{\mathbf{v}}\exp(i\mathbf{vx})] \quad . \quad (4.7)$$

will now be treated as perturbation. This can be done by straightforward use of perturbation theory provided the zero order energy of the system is less than $\hbar\omega$, i.e. the corresponding eigenvalue ϵ_0 of F_0 satisfies $\epsilon_0<1$, because then for given total wave vector $\mathbf{k}$ the system is not degenerate. For in this case no quantum is excited, hence $\mathbf{k}=\mathbf{q}$, and $k<1$ the latter to satisfy $\epsilon_0<1$. Thus

$$|\,\mathbf{k}\,;\ldots 0_{\mathbf{v}}\ldots\rangle=S^{-1/2}\exp(i\mathbf{kx})\chi\,;\quad \chi=\prod_{\mathbf{v}}|\,0_{\mathbf{v}}\rangle\,;\quad k<1 \quad . \quad (4.8)$$

is the normalized zero order eigenfunction with wave vector $\mathbf{k}$, and

$$\epsilon_0(\mathbf{k},\ldots 0_{\mathbf{v}}\ldots)=k^2<1. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.9)$$

is the corresponding zero order energy.

Now since according to (3.5) and (3.6)

$$b_{\mathbf{v}}|\,0_{\mathbf{v}}\rangle=0\,;\quad b_{\mathbf{v}}^{+}|\,0_{\mathbf{v}}\rangle=|\,1_{\mathbf{v}}\rangle, \quad . \quad . \quad . \quad . \quad (4.10)$$

and in view of the orthogonality of the wave functions

$$S^{-1}\int_S\exp(i\mathbf{k}-\mathbf{k}'\,.\,\mathbf{x})\,d^3\mathbf{k}=\delta_{\mathbf{kk}'}\,;\quad \langle n_{\mathbf{v}}|\,n_{\mathbf{v}'}\rangle=\delta_{\mathbf{vv}'}, \quad . \quad (4.11)$$

the interaction term F_{int} leads to transitions from a state described by (4.8) to states in which a single quantum is excited, provided the total wave vector remains unaltered. Their wave functions are thus of the type

$$|\,\mathbf{k}\,;\,0_{\mathbf{v}_1}\ldots 1_{\mathbf{v}}\ldots 0_{\mathbf{v}'}\ldots\rangle=S^{-1/2}\exp(i\mathbf{k}-\mathbf{v}\,.\,\mathbf{x})b_{\mathbf{v}}^{+}\chi, \quad . \quad (4.12)$$

with an energy

$$\epsilon_0(\mathbf{k}\,;\,0_{\mathbf{v}_1}\ldots 1_{\mathbf{v}}\ldots)=(\mathbf{k}-\mathbf{v})^2+1. \quad . \quad . \quad . \quad . \quad (4.13)$$

The non vanishing matrix elements are thus given by

$$\langle\ldots 1_{\mathbf{v}}\ldots 0_{\mathbf{v}_1}\,;\,\mathbf{k}|\,F_{\text{int}}|\,\mathbf{k}\,;\,0_{\mathbf{v}_1}\ldots 1_{\mathbf{v}}\ldots\rangle=i(4\pi\alpha S)^{1/2}1/v. \quad . \quad (4.14)$$

From second order perturbation theory, therefore,

$$E(\mathbf{k})=\hbar\omega\left(k^2-\frac{4\pi\alpha}{S}\sum_{\mathbf{v}}\frac{1}{v^2[(\mathbf{k}-\mathbf{v})^2+1-k^2]}\right) \quad . \quad . \quad (4.15)$$

is the perturbed energy (in ordinary units, hence the factor $\hbar\omega$), and. using (4.8), (4.9), (4.12)–(4.14),

$$\Psi(\mathbf{k})=\frac{\exp(i\mathbf{kx})}{S^{1/2}}\left(1-i\left(\frac{4\pi\alpha}{S}\right)^{1/2}\sum_{\mathbf{v}}\frac{b_{\mathbf{v}}^{+}\exp(-i\mathbf{vx})}{v[(\mathbf{k}-\mathbf{v})^2+1-k^2]}\right)\chi, \quad . \quad (4.16)$$

the corresponding wave function.

The summation in (4.15) is easily carried out by transforming the sum into an integral,

$$\sum_{\mathbf{v}}\to\frac{S}{(2\pi)^3}\int\ldots d^3\mathbf{v}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.17)$$

which follows from (2.29) and (2.39). It will be observed that no restriction in the **v**-integration limits of the type mentioned in § 2 for the classical case is required because of the rapid convergence of the integral for $v>1$. Thus

$$E(\mathbf{k})=\hbar\omega\left(k^2-\alpha\frac{\sin^{-1}k}{k}\right)\;;\quad k<1,\ \alpha<1. \quad . \quad . \quad . \quad (4.18)$$

Also since for $k<1$,

$$\frac{\sin^{-1}k}{k}=k+\tfrac{1}{6}k^3+\ldots,\quad k<1 \quad . \quad . \quad . \quad . \quad (4.19)$$

we have

$$E(\mathbf{k})=E(0)+\frac{\hbar^2}{2m^*}(uk)^2+\ldots, \quad . \quad . \quad . \quad . \quad . \quad (4.20)$$

where (c.f (2.38))

$$E(0)=-\alpha\hbar\omega=-\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right)e^2\left(\frac{m\omega}{2\hbar}\right)^{1/2},\quad \alpha<1 \quad . \quad . \quad (4.21)$$

is the energy of the lowest state ($k=0$) and

$$m^*=m/(1-\alpha/6)\simeq m(1+\alpha/6),\qquad \alpha<1 \quad . \quad . \quad . \quad (4.22)$$

may be denoted as an effective mass. For uk is the wave number in usual units, and hence $(\hbar uk)^2/2m^*$ would represent the kinetic energy of a particle of mass m^* and momentum $\hbar uk$.

The perturbation method described above is restricted to weak interaction, $\alpha<1$. It leads to wave functions of type (4.16) in which never more than a single quantum can be excited. It will be shown now that the range of validity of eqn. (4.21) goes far beyond $\alpha=1$.

4.2. *Gurari Method*

Gurari has shown (1953) that a variational method can be applied in order to derive eqn. (4.21) for $E(0)$. In § 5 with a different variational method another value for $E(0)$ will be derived which is proportional to α^2. This will permit us to estimate the α-range in which (4.21) holds with reasonable accuracy because variational methods never yield energies lower than the exact one. Gurari's results were derived independently by Lee and Pines (1952) who use a method by Tomonaga (1947) which is closely related to Gurari's method (cf. also Lee, Low and Pines 1953, and Zienau 1953).

In the general theory of quantum mechanics it is shown that the solutions of

$$F\Psi=E(\mathbf{k})\Psi, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.23)$$

satisfy the variational principle

$$\delta J=0 \qquad \text{with} \qquad \int\Psi^+\Psi=1, \quad . \quad . \quad . \quad . \quad (4.24)$$

where

$$J=\int\Psi^+F\Psi. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.25)$$

The integrations are to be extended over all variables. They are understood in a sufficiently general sense, e.g. if the $b_{\mathbf{w}}$'s are represented by matrices then the corresponding sums must be taken.

The functions Ψ will now be chosen through a generalization of the wave functions in perturbation theory. Instead of using a linear superposition of wave functions (all with equal total wave vector $\mathbf{k}$) containing no more than one quantum, $\Psi(\mathbf{k})$ will now be chosen so as to lead to no more than a single quantum for each wave vector $\mathbf{v}$; any total number of quanta may be present, however, provided they belong to different wave vectors. Again let $\mathbf{k}$ be the total wave vector of the system, and define functions

$$\psi(\mathbf{k}, \mathbf{v})=N(k, v, \xi)[1+S^{-1/2}c(k, v, \xi) \exp(-i\mathbf{vx})b_{\mathbf{v}}^{+}], \quad . \quad (4.26)$$

where

$$N(k, v, \xi)=[1+S^{-1} \mid c(k, v, \xi) \mid^2]^{-1/2} \simeq 1. \quad . \quad . \quad . \quad (4.27)$$

The $c(k, v, \xi)$ are numerical constants, and ξ is the cosine of the angle between $\mathbf{k}$ and $\mathbf{v}$,

$$\xi=(\mathbf{kv})/kv. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.28)$$

The volume S (in reduced units) is supposed to be extended towards infinity—after completion of the calculation of $E(\mathbf{k})$; its occurrence thus defines orders of magnitudes and it will be assumed that the c's do not contain S. Let then

$$\Psi(\mathbf{k})=S^{-1/2} \exp(i\mathbf{kx}) \prod_{\mathbf{v}} \psi(\mathbf{k}, \mathbf{v})\chi, \quad . \quad . \quad . \quad . \quad (4.29)$$

and determine the c's from the variational eqn. (4.24). It will be noticed that $\Psi(\mathbf{k})$ is already normalized. Also by development of the product into a sum it is easily seen that all terms occurring are of the form

$$\mid \mathbf{k}; \ldots n_{\mathbf{v}} \ldots \rangle = S^{-1/2} \exp[i(\mathbf{k}-\Sigma \mathbf{v}n_{\mathbf{v}})\mathbf{x}] \prod_{\mathbf{v}} \mid n_{\mathbf{v}} \rangle, \quad . \quad (4.30)$$

thus, according to (4.2) and (4.3), belonging to the same total wave vector $\mathbf{k}$.

In calculating the integral J, (4.25) with the use of (2.40) and (4.29) consider first the pure field terms $\Sigma_{\mathbf{v}} b_{\mathbf{v}}^{+}b_{\mathbf{v}}$. It will be noticed that all $\psi(\mathbf{k}, \mathbf{v})$ commute, and that $b_{\mathbf{v}}^{+}b_{\mathbf{v}}$ commutes with all $\psi(\mathbf{k}, \mathbf{v}')$'s $(\mathbf{v} \neq \mathbf{v}')$. Hence using χ from (4.8)

$$J_1 = \int \Psi^{+} \sum_{\mathbf{v}} b_{\mathbf{v}}^{+}b_{\mathbf{v}}\Psi = S^{-1} \sum_{\mathbf{v}} \int_S d^3\mathbf{x} \langle 0_{\mathbf{v}} \mid \psi^{+}(\mathbf{k}, \mathbf{v})b_{\mathbf{v}}^{+}b_{\mathbf{v}}\psi(\mathbf{k}, \mathbf{v}) \mid 0_{\mathbf{v}} \rangle$$

$$= S^{-1} \sum_{\mathbf{v}} \int_S N^2(v, k, \xi) \mid c(k, v, \xi) \mid^2 S^{-1} d^3\mathbf{x} \simeq S^{-1} \sum_{\mathbf{v}} \mid c(k, v, \xi) \mid^2. \quad . \quad (4.31)$$

Here use has been made of (4.10) and (4.11), and $N^2 \simeq 1$ has been introduced because $S \to \infty$. It must be observed that this is always possible in sums of the type occurring in (4.31) and in similar ones occurring below. In the wave function $\Psi(\mathbf{k})$, however, where a product $\Pi_{\mathbf{v}} N(v, k, \xi)$ occurs in the denominator replacement of $N(v, k, \xi)$ by unity is not permissible because the number of factors in the product increases with S.

The interaction term in F reduces in a similar way to a sum when eqns. (4.10) and (4.11) are used. For then

$$J_2=i(4\pi\alpha/S)^{1/2}\int \Psi^+ \sum_{\mathbf{v}} [b_{\mathbf{v}}^+ \exp(-i\mathbf{vx})-b_{\mathbf{v}}\exp(i\mathbf{vx})]\Psi$$

$$=iS^{-3/2}(4\pi\alpha)^{1/2}\sum_{\mathbf{v}}\int_S d^3\mathbf{x}\,\langle 0_{\mathbf{v}}\,|\,\psi^+(\mathbf{k},\mathbf{v})[b_{\mathbf{v}}^+\exp(-i\mathbf{vx})$$

$$-b_{\mathbf{v}}\exp(i\mathbf{vx})]\psi(\mathbf{k},\mathbf{v})\,|\,0_{\mathbf{v}}\rangle$$

$$=iS^{-1}(4\pi\alpha)^{1/2}\sum_{\mathbf{v}}[c^*(k,v,\xi)-c(k,v,\xi)]/v. \quad . \quad . \quad . \quad . \quad . \quad (4.32)$$

Finally the electronic term is slightly more complicated,

$$J_3=-\sum_1^3\int \Psi^+\frac{\partial^2}{\partial x_i^{\,2}}\Psi=\int\frac{\partial\Psi^+}{\partial\mathbf{x}}\frac{\partial\Psi}{\partial\mathbf{x}}=k^2-iS^{-1}\mathbf{k}\int\chi^+\prod_{\mathbf{v}}\psi^+(\mathbf{k},\mathbf{v})$$

$$\times\frac{\partial}{\partial\mathbf{x}}\prod_{\mathbf{v}}\psi(\mathbf{k},\mathbf{v})\chi+\text{c.c.}+S^{-1}\int\chi^+\left(\frac{\partial}{\partial\mathbf{x}}\prod_{\mathbf{v}}\psi^+(\mathbf{k},\mathbf{v})\right)\left(\frac{\partial}{\partial\mathbf{x}}\prod_{\mathbf{v}}\psi(\mathbf{k},\mathbf{v})\right)\chi.$$

$$. \quad . \quad . \quad . \quad (4.33)$$

Here c.c. denotes the conjugate complex of the previous term. The second term in (4.33) is obtained by similar methods to the ones used to evaluate (4.31) and (4.32). Thus with χ from (4.8) and with use of (4.11),

$$-iS^{-1}\mathbf{k}\int\chi^+\prod_{\mathbf{v}}\psi^+(\mathbf{k},\mathbf{v})\frac{\partial}{\partial\mathbf{x}}\prod_{\mathbf{v}}\psi(\mathbf{k},\mathbf{v})\chi$$

$$=-iS^{-1}\mathbf{k}\sum_{\mathbf{v}}\int_S d^3\mathbf{x}\,\langle 0_{\mathbf{v}}\,|\,\psi^+(\mathbf{k},\mathbf{v})\frac{\partial}{\partial\mathbf{x}}\psi(\mathbf{k},\mathbf{v})\,|\,0_{\mathbf{v}}\rangle$$

$$=-iS^{-1}\mathbf{k}\sum_{\mathbf{v}}\mathbf{v}\,|\,c(k,v,\xi)\,|^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.34)$$

In a similar way the last term is found, and one obtains

$$J_3=k^2-2S^{-1}\sum_{\mathbf{v}}(\mathbf{kv})\,|\,c(k,v,\xi)\,|^2+S^{-1}\sum_{\mathbf{v}}v^2\,|\,c(k,v,\xi)\,|^2$$

$$+S^{-2}\sum_{\mathbf{v}}\sum_{\mathbf{v}'}(\mathbf{vv}')\,|\,c(k,v,\xi)\,|^2\,|\,c(k,v',\xi')\,|^2. \quad . \quad . \quad . \quad (4.35)$$

Here the factor $(\mathbf{vv}')$ can be simplified because using (4.28)

$$\mathbf{vv}'=vv'(\xi\xi'+\sin(\mathbf{k},\mathbf{v})\sin(\mathbf{k},\mathbf{v}')\cos(\phi-\phi')) \quad . \quad . \quad (4.36)$$

where ϕ and ϕ' are the polar angles of $\mathbf{v}$ and $\mathbf{v}'$ respectively. In the following summation (integration) over $\mathbf{v}$, the term with $\cos(\phi-\phi')$ vanishes because $c(k,v,\xi)$ is independent of ϕ, and

$$\int_0^{2\pi}\cos(\phi-\phi')\,d\phi=0.$$

Thus $\mathbf{vv}'$ can be replaced by $vv'\xi\xi'$. Hence using (4.31), (4.32) and (4.35),

$$J=J_1+J_2+J_3=k^2+S^{-1}\sum_{\mathbf{v}}(1-2\mathbf{kv}+v^2)c^*(k,v,\xi)c(k,v,\xi)$$

$$+i(4\pi\alpha)^{1/2}S^{-1}\sum_{\mathbf{v}}(c^*(k,v,\xi)-c(k,v,\xi))/v+g^2 \quad . \quad . \quad . \quad . \quad (4.37)$$

where

$$g=S^{-1}\sum_{\mathbf{v}} v\xi c^*(k, v, \xi)c(k, v, \xi). \quad . \quad . \quad . \quad . \quad (4.38)$$

The variational principle $\delta J=0$ requires that differentiation of J with regard to all $c(k, v, \xi)$ and $c^*(k, v, \xi)$ should vanish. Hence

$$0=(1-2kv\xi+2gv\xi+v^2)c(k, v, \xi)+i(4\pi\alpha)^{1/2}/v. \quad . \quad . \quad (4.39)$$

It is useful to introduce here a vector $\mathbf{g}$ whose magnitude is g and whose direction is the same as that of $\mathbf{k}$, i.e.

$$\mathbf{g}=g\mathbf{k}/k. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.40)$$

Then from (4.39)

$$c(k, v, \xi)=-\frac{i(4\pi\alpha)^{1/2}}{v(1-2(\mathbf{k}-\mathbf{g}\,.\,\mathbf{v})+v^2)}, \quad . \quad . \quad . \quad (4.41)$$

and $c^*(k, v, \xi)$ is its conjugate complex.

The above solution holds only if $c(k, v, \xi)$ does not become infinite for and value of $\mathbf{v}$. To prevent this the condition

$$(k-g)^2<1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.42)$$

must be satisfied.

With the values (4.41) for the coefficients c the integral J, (4.25), represents the energy (in terms of $\hbar\omega$). Hence from (4.37) and (4.41), using (4.38) and (4.40),

$$\begin{aligned} E(\mathbf{k})&=\hbar\omega\left\{k^2+g^2+4\pi\alpha S^{-1}\sum_{\mathbf{v}}\left(\frac{1-2\mathbf{kv}+v^2}{v^2[1-2(\mathbf{k}-\mathbf{g}\,.\,\mathbf{v})+v^2]^2}\right.\right. \\ &\qquad\left.\left.-\frac{2}{v^2[1-2(\mathbf{k}-\mathbf{g}\,.\,\mathbf{v})+v^2]}\right)\right\} \\ &=\hbar\omega\left\{k^2-g^2-4\pi\alpha S^{-1}\sum_{\mathbf{v}}\frac{1}{v^2[1-2(\mathbf{k}-\mathbf{g}\,.\,\mathbf{v})+v^2]}\right\} \quad . \quad (4.43) \end{aligned}$$

and from (4.38) and (4.41), using (4.40),

$$g=\frac{4\pi\alpha}{k}S^{-1}\sum_{\mathbf{v}}\frac{(\mathbf{vk})}{v^2[1-2(\mathbf{k}-\mathbf{g}\,.\,\mathbf{v})+v^2]^2}. \quad . \quad . \quad . \quad (4.44)$$

With (4.17) after integration one obtains

$$E(\mathbf{k})=\hbar\omega\left\{k^2-g^2-\alpha\frac{\sin^{-1}(k-g)}{k-g}\right\}, \quad . \quad . \quad . \quad . \quad (4.45)$$

and

$$g=\tfrac{1}{2}\alpha\left\{-\frac{\sin^{-1}(k-g)}{(k-g)^2}+\frac{1}{(k-g)(1-(k-g)^2)^{1/2}}\right\}, \quad . \quad (4.46)$$

i.e. two simultaneous equations for $E(k)$ and g. Again condition (4.42) must be observed. It will be noticed that eqn. (4.46) can also be written as

$$g=\tfrac{1}{2}\alpha\frac{d}{dk}\frac{\sin^{-1}(k-g)}{k-g}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.47)$$

Now if $|k-g| \ll 1$, by development of $[\sin^{-1}(k-g)]/(k-g)$ (cf. (4.19)), eqn. (4.47) becomes

$$g=\alpha(k-g)/6, \quad \text{or} \quad k=(k-g)(1+\alpha/6), \quad \text{or} \quad g=\frac{k\alpha/6}{1+\alpha/6}. \quad . \quad . \quad . \quad (4.48)$$

Thus as $k \to 0$, also $g \to 0$. With the use of (4.48), and of a further development of the $\sin^{-1}$, finally

$$E(\mathbf{k})=E(0)+\hbar^2(ku)^2/2m^*, \qquad k \ll 1, \quad . \quad . \quad . \quad . \quad . \quad (4.49)$$

where

$$E(0)=-\alpha\hbar\omega=-\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right)e^2\left(\frac{m\omega}{2\hbar}\right)^{1/2}, \quad . \quad . \quad (4.50)$$

and

$$m^*=m(1+\alpha/6). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.51)$$

Thus $E(0)$ is identical with the result (4.21) of perturbation theory, but it is no longer restricted to $\alpha<1$. The limit of α up to which the result is reliable will be studied in § 5 with the use of a variational method which gives the energy proportional to α^2. Also m^* is seen to agree with the result which perturbation theory yields for $\alpha \ll 1$. Again this value now holds for larger α.

Expression (4.49) had to be restricted to sufficiently small values of k to make use of the development of $\sin^{-1}(k-g)$. This would not restrict, however, the validity of eqns. (4.45) and (4.46) which have solutions for all k. In particular $E(k) \propto k\hbar\omega$ as $k \to \infty$. Such solutions, however, have no physical meaning because $E(\mathbf{k})$ has to be restricted to values for which

$$E(\mathbf{k})-E(0)<\hbar\omega \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.52)$$

which means that the energy of the system should be less than one quantum above the energy of the ground state. If this were not fulfilled then the trial wave functions (4.29) (with (4.26)) would not be the only ones leading to the energy $E(\mathbf{k})$ for given $\mathbf{k}$. Others would exist as well in which at least one of the $\psi(\mathbf{k}, \mathbf{v})$'s would be replaced by a function in which the state with one quantum is principally excited (and not the state without quantum as in (4.29)). A degeneracy would thus occur. It would require a new approach which has not been found yet.

Equation (4.49) suggests that the electric current carried by the system is given by $ej=e\hbar ku/m^*$ where j represents the average electronic velocity. In fact with the use of (4.29), (4.26), (4.41), (4.48) and of the orthogonality relations,

$$\begin{aligned} ej &= \frac{e\hbar u}{m}\int \Psi^+\frac{\partial\Psi}{\partial\mathbf{x}} = \frac{e\hbar}{m}\left\{\mathbf{k}-4\pi\alpha S^{-1}\sum_{\mathbf{v}}\frac{\mathbf{v}}{v^2(1-2(\mathbf{k}-\mathbf{g}\,.\,\mathbf{v})+v^2)^2}\right\} \\ &= \frac{e\hbar u}{m}\frac{\mathbf{k}}{k}(k-g)=\frac{e\hbar u}{m^*}\mathbf{k}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.53) \end{aligned}$$

as suggested. Use has been made of (4.44) and of the fact that the direction of the vector represented by the sum must be $\mathbf{k}/k$. In the last step (4.51) has been applied.

It is also of interest to calculate the polarization near an electron. For this purpose it is best to derive an appropriate expectation value of the polarization potential $\Phi_{ir}(\mathbf{r})$ defined in (2.30). Clearly with (2.25) and (2.28), using (2.37)–(2.39)

$$e\Phi_{ir}(\mathbf{y})=i(4\pi\alpha/S)^{1/2}\hbar\omega\sum_{\mathbf{v}}(b_{\mathbf{v}}^{+}\exp(-i\mathbf{vy})-b_{\mathbf{v}}\exp(i\mathbf{vy}))/v, \quad . \quad (4.54)$$

where

$$\mathbf{y}=u\mathbf{r} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.55)$$

has been introduced. Let now the integration over all variables of the wave function Ψ be split into an electronic one, denoted by $\int d^3\mathbf{x}$, and one, denoted by the symbol $\langle \quad \rangle$, referring to the field variables. Thus the integration over all variables introduced in (4.25) can be written as

$$\int=\int d^3\mathbf{x}\langle \quad \rangle. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.56)$$

It follows then that $\langle\Psi^{+}\Psi\rangle$ is the probability of finding the electron at a position $\mathbf{x}$ irrespective of the polarization, and $\langle\Psi^{+}\Phi_{ir}(\mathbf{y})\Psi\rangle$ represents the expectation value of $\Phi_{ir}(\mathbf{y})$ at a position $\mathbf{y}$ and simultaneously finding the electron at $\mathbf{x}$.

Define now

$$\Phi_{ir}(\mathbf{x},\mathbf{y})=\langle\Psi^{+}\Phi_{ir}(\mathbf{y})\Psi\rangle/\langle\Psi^{+}\Psi\rangle=S\langle\Psi^{+}\Phi_{ir}(\mathbf{y})\Psi\rangle, \quad . \quad (4.57)$$

which is the expectation value of the polarization potential at a position $\mathbf{y}$, provided the electron is at $\mathbf{x}$. For this purpose it will be noticed that from (4.29) and (4.26), with the use of (3.5), (4.10) and (4.11),

$$\begin{aligned}\langle\Psi^{+}b_{\mathbf{v}'}^{+}\Psi\rangle&=S^{-1}\langle\chi^{+}\prod_{\mathbf{v}}\psi^{+}(\mathbf{k},\mathbf{v})b_{\mathbf{v}'}^{+}\prod_{\mathbf{v}}\psi(\mathbf{k},\mathbf{v})\chi\rangle\\&=S^{-1}\langle 0_{\mathbf{v}'}[1+S^{-1/2}c^{*}(k,v',\xi')\exp(i\mathbf{v}'\mathbf{x})b_{\mathbf{v}'}]\\&\quad\times b_{\mathbf{v}'}^{+}[1+S^{-1/2}c(k,v',\xi')\exp(-i\mathbf{v}'\mathbf{x})b_{\mathbf{v}'}^{+}]0_{\mathbf{v}'}\rangle\\&=S^{-3/2}c^{*}(k,v',\xi')\exp(i\mathbf{v}'\mathbf{x}). \quad . \quad . \quad . \quad . \quad . \quad (4.58)\end{aligned}$$

Hence from (4.54), (4.57) and (4.41)

$$\begin{aligned}e\Phi_{ir}(\mathbf{x},\mathbf{y})&=i(4\pi\alpha)^{1/2}S^{-1}\hbar\omega\sum_{\mathbf{v}}[c^{*}(k,v,\xi)\exp(i\mathbf{v}\,.\,\mathbf{x}-\mathbf{y})\\&\quad-c(k,v,\xi)\exp(-i\mathbf{v}\,.\,\mathbf{x}-\mathbf{y})]/v\\&=-4\pi\alpha S^{-1}\sum_{\mathbf{v}}\frac{2\cos(\mathbf{v}\,.\,\mathbf{k}-\mathbf{y})}{v^2[1-2(\mathbf{k}-\mathbf{g}\,.\,\mathbf{v})+v^2]}. \quad . \quad . \quad . \quad (4.59)\end{aligned}$$

In the simplest case $k=0$ (i.e. also $g=0$) one has with (4.17),

$$\begin{aligned}S^{-1}\sum_{\mathbf{v}}\frac{2\cos(\mathbf{v}\,.\,\mathbf{x}-\mathbf{y})}{v^2(1+v^2)}&=\frac{2}{4\pi^2}\int_0^{\infty}dv\int_{-1}^{1}\frac{\cos(v|\mathbf{x}-\mathbf{y}|\eta)}{1+v^2}d\eta\\&=\frac{1}{\pi^2|\mathbf{x}-\mathbf{y}|}\int_0^{\infty}\frac{\sin(v|\mathbf{x}-\mathbf{y}|)}{v(1+v^2)}dv\\&=\frac{1-\exp(-|\mathbf{x}-\mathbf{y}|)}{\pi|\mathbf{x}-\mathbf{y}|}. \quad . \quad . \quad . \quad . \quad (4.60)\end{aligned}$$

Returning now to ordinary units with the help of (2.37)–(2.39) (denoting then $\Phi_{ir}(\mathbf{x}, \mathbf{y})$ by $\Phi_{ir}(\mathbf{r}_{el}, \mathbf{r})$) eqns. (4.59) and (4.60) yield

$$e\Phi_{ir}(\mathbf{r}_{el}, \mathbf{r}) = -\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon}\right) \frac{e^2}{|\mathbf{r}-\mathbf{r}_{el}|} [1 - \exp(-u\,|\,\mathbf{r}-\mathbf{r}_{el}\,|)]. \qquad (4.61)$$

Thus if $|\mathbf{r}-\mathbf{r}_{el}| \gg 1\,u$, the polarization potential of a point charge is found. If $|\mathbf{r}-\mathbf{r}_{el}| \ll 1/u$, however, one finds

$$e\Phi_{ir}(\mathbf{r}_{el}, \mathbf{r}) \simeq -\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon}\right) e^2 u, \qquad |\mathbf{r}-\mathbf{r}_{el}| \ll 1/u, \qquad (4.62)$$

i.e. the polarization potential of a charge extended over a sphere whose radius is of the order $1/u$. The case that k is different from zero leads to a slight modification of (4.61).

§ 5. Stationary Solutions II

The eigenvalues of the Hamiltonian (2.40) derived in the previous section are linear in the interaction parameter α. In the present section a different method, originally due to Landau (1933) (cf. also Pekar 1946, 1949) will be used which has been discussed qualitatively in § 1. It will be shown that it can be derived with the help of the variational principle (4.24) and will be found to yield eigenvalues which are quadratic in α.

For this purpose use the trial wave function Ψ as a product of an electronic wave function $\Omega(\mathbf{x})$ and a lattice wave function Φ,

$$\Psi = \Omega(\mathbf{x})\Phi, \qquad (5.1)$$

where $\Omega(\mathbf{x})$ depends on the electronic coordinate $\mathbf{x}$ only, and Φ on the coordinates describing the polarization waves. It will be assumed that both are normalized, i.e. using the definition (4.56) for the respective integrations

$$\int_S d^3\mathbf{x}\,\Omega^+(\mathbf{x})\Omega(\mathbf{x}) = 1\,; \qquad \langle \Phi^+\Phi \rangle = 1. \qquad (5.2)$$

Clearly with (4.1) and (4.56),

$$\int \Psi^+ F_0 \Psi = -\sum_1^3 \int_S \Omega^+(\mathbf{x}) \frac{\partial^2}{\partial x_i^2} \Omega(\mathbf{x})\, d^3\mathbf{x} + \sum_{\mathbf{v}} \langle \Phi^+ b_{\mathbf{v}}^+ b_{\mathbf{v}} \Phi \rangle, \qquad (5.3)$$

and with (4.7)

$$\int \Psi^+ F_{\text{int}} \Psi = i(4\pi\alpha/S)^{1/2} \langle \Phi^+ \int_S d^3\mathbf{x}\,\Omega^+(\mathbf{x}) \sum_{\mathbf{v}} [b_{\mathbf{v}}^+ \exp(-i\mathbf{v}\mathbf{x}) - b_{\mathbf{v}} \exp(i\mathbf{v}\mathbf{x})]\Omega(\mathbf{x})\Phi \rangle. \qquad (5.4)$$

The conditions (5.2) can be satisfied with the use of the Lagrangian factors λ_1 and λ_2 so that the variational principle (4.24) requires

$$\delta\left\{\int \Psi^+ F \Psi - \lambda_1 \int_S \Omega^+\Omega\, d^3\mathbf{k} - \lambda_2 \langle \Phi^+\Phi \rangle\right\} = 0, \qquad (5.5)$$

or using (5.3) and (5.4),

$$\delta\left\{-\int_S d^3\mathbf{x}\,\Omega^+(\mathbf{x})\left(\sum_1^3\frac{\partial^2}{\partial x_i^{\,2}}+\lambda_1\right)\Omega(\mathbf{x})+\sum_{\nu}\langle\Phi^+(\Sigma b_\nu^+ b_\nu-\lambda_2)\Phi\rangle\right.$$
$$\left.+\langle\Phi^+\int_S d^3\mathbf{x}\,\Omega^+(\mathbf{x})F_{\text{int}}\Omega(\mathbf{x})\Phi\rangle\right\}=0, \quad . \quad . \quad . \quad . \quad (5.6)$$

where the functions Ω, Φ and their conjugates have to be varied. This is simplified because only the last term depends on both. Thus variation with respect to $\Omega^+(\mathbf{x})$ yields

$$-\sum_1^3\frac{\partial^2\Omega(\mathbf{x})}{\partial x_i^{\,2}}+\langle\Phi^+F_{\text{int}}\Phi\rangle\Omega(\mathbf{x})=\lambda_1\Omega(\mathbf{x}^*). \quad . \quad . \quad . \quad (5.7)$$

Similarly from variation of Φ^+,

$$\Sigma b_\nu^+ b_\nu\Phi+\int_S d^3\mathbf{x}\,\Omega^+(\mathbf{x})F_{\text{int}}\Omega(\mathbf{x})\Phi=\lambda_2\Phi. \quad . \quad . \quad . \quad (5.8)$$

Variation of Ω and Φ leads to the conjugate equations. Equation (5.7) is the Schrödinger equation for an electron moving in a ' self consistent ' field $\langle\Phi^+F_{\text{int}}\Phi\rangle$. Similarly (5.8) represents the motion of the field of polarization waves in a ' self consistent ' source. Clearly in view of (5.4), Φ can be written as a product,

$$\Phi=\Pi\Phi_\nu; \quad \text{also } \lambda_2=\Sigma\lambda_\nu, \quad . \quad . \quad . \quad . \quad . \quad (5.9)$$

where Φ_ν depends on the coordinates of the ν-th field component only. It satisfies

$$b_\nu^+ b_\nu\Phi_\nu+i(4\pi\alpha/S)^{1/2}(1/v)(b_\nu^+\int_S\Omega^+(\mathbf{x})\exp(-i\nu\mathbf{x})\Omega(\mathbf{x})\,d^3\mathbf{x}$$
$$-b_\nu\int_S\Omega^+(\mathbf{x})\exp(i\nu\mathbf{x})\Omega(\mathbf{x})\,d^3\mathbf{x})\Phi_\nu=\lambda_\nu\Phi_\nu, \quad . \quad . \quad (5.10)$$

which is the equation of a displaced harmonic oscillator. Hence

$$\langle\Phi_\nu^+ b_\nu^+\Phi_\nu\rangle=i(4\pi\alpha/S)^{1/2}(1/v)\int_S\Omega^+(\mathbf{x})\exp(i\nu\mathbf{x})\Omega(\mathbf{x})\,d^3\mathbf{x}. \quad . \quad (5.11)$$

With the use of (5.4) and (5.11), therefore, the self consistent potential energy of the electron becomes (use (4.17))

$$\langle\Phi^+F_{\text{int}}\Phi\rangle$$
$$=-\frac{4\pi\alpha}{S}\sum_\nu\frac{1}{v^2}\int_S d^3\mathbf{x}'\,\Omega^+(\mathbf{x}')\Omega(\mathbf{x}')[\exp(i\nu\,.\,\mathbf{x}'-\mathbf{x})+\exp(-i\nu\,.\,\mathbf{x}'-\mathbf{x})]$$
$$=\frac{4\pi\alpha}{(2\pi)^3}\int_S d^3\mathbf{x}'\int d^3\nu\,\frac{2\cos(\nu\,.\,\mathbf{k}'-\mathbf{x})}{v^2}\,\Omega^+(\mathbf{x}')\Omega(\mathbf{x}')$$
$$=-2\alpha\int_S\frac{\Omega^+(\mathbf{x}')\Omega(\mathbf{x}')}{|\mathbf{x}'-\mathbf{x}|}\,d^3\mathbf{x}'. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.12)$$

It represents the energy of the charge e at $\mathbf{x}$ placed into the polarization potential due to the charge distribution $\Omega^+(\mathbf{x}')\Omega(\mathbf{x}')$. This is easily seen by going back to conventional units with the help of (2.37)–(2.39), i.e.

$$\hbar\omega\langle\Phi^+F_{\text{int}}\Phi\rangle=-\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right)e^2\int\frac{\Omega^+(\mathbf{r})\Omega(\mathbf{r})}{|\mathbf{r}-\mathbf{r}_{el}|}d^3\mathbf{r}. \quad . \quad . \quad (5.13)$$

Introduction of (5.13) into (5.7) leads to an integral equation for $\Omega(\mathbf{x})$. Qualitatively, it is clear that $\Omega(\mathbf{x})$ must be a function which is large in a limited space only. For only in this way can the ' self consistent ' potential energy have relatively large negative values. To find an approximative solution use will be made again of the variational method in its original form (4.24). For this purpose let

$$\Omega(\mathbf{x})=(\beta^3/8\pi)^{1/2}\exp(-\beta|\mathbf{x}-\mathbf{x}_0|/2), \quad . \quad . \quad . \quad (5.14)$$

where $\mathbf{x}_0$ is an arbitrary radius vector. Then with

$$J=\langle\int_S d^3\mathbf{x}\Phi^+\Omega^+(\mathbf{x})F\Omega(\mathbf{x})\Phi\rangle, \quad . \quad . \quad . \quad . \quad (5.15)$$

the variational method requires that

$$\partial J/\partial\beta=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.16)$$

provided Φ and Ω are normalized, and Φ is a solution of (5.8). As mentioned it is composed of displaced harmonic oscillator solutions.

Now using the expression for F according to (5.3) and (5.4), together with (5.8)–(5.10), the integral J, defined in (5.15) becomes

$$J=-\int_S\Omega^+(\mathbf{x})\sum_1^3\frac{\partial^2}{\partial x_i^2}\Omega(\mathbf{x})\,d^3\mathbf{x}+\sum_\nu\lambda_\nu. \quad . \quad . \quad (5.17)$$

Also from (5.10) for the lowest states of the displaced harmonic oscillators,

$$\begin{aligned}\sum_\nu\lambda_\nu&=-(4\pi\alpha/S)\sum_\nu\left|\int_S\Omega^+(\mathbf{x})\exp(-i\nu\mathbf{x})\Omega(\mathbf{x})\,d^3\mathbf{x}\right|^2\Big/v^2\\&=-(4\pi\alpha/S)\int_S d^3\mathbf{x}\Big\{\Omega^+(\mathbf{x})\Omega(\mathbf{x})\tfrac{1}{2}\sum_\nu(1/v^2)\int_S d^3\mathbf{x}'\,\Omega^+(\mathbf{x}')\Omega(\mathbf{x}')\\&\quad\times[\exp(i\nu\,.\,\mathbf{x}'-\mathbf{x})+\exp(-i\nu\mathbf{x}'-\mathbf{x})]\Big\}\\&=\frac{1}{2}\int_S\Omega^+(\mathbf{x})\Omega(\mathbf{x})\langle\Phi^+F_{\text{int}}\Phi\rangle\,d^3\mathbf{x}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.18)\end{aligned}$$

where eqn. (5.12) has been used in the last step. Thus the potential energy $\sum\lambda_\nu$ appearing in (5.17) is obtained by averaging the potential energy (5.12) over the density distribution $\Omega^+(\mathbf{x})\Omega(\mathbf{x})$ of the electron, and multiplying it by $\frac{1}{2}$. This reduction by a factor $\frac{1}{2}$ arises from the positive contributions in displacing the oscillators.

Introducing now $\Omega(\mathbf{x})$ from (5.14), one finds with (5.18) and (5.12),

$$\begin{aligned}\sum_\nu\lambda_\nu&=-\alpha(\beta^3/8\pi)^2\int_S d^3\mathbf{x}\int_S d^3\mathbf{x}'\,\frac{\exp(-\beta|\mathbf{x}'-\mathbf{x}_0|)\exp(-\beta|\mathbf{x}-\mathbf{x}_0|)}{|\mathbf{x}-\mathbf{x}'|}\\&=-\frac{5}{16}\alpha\beta. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.19)\end{aligned}$$

Furthermore,

$$-\int_S \Omega^+(\mathbf{x})\sum_1^3 \frac{\partial^2}{\partial x_i^2}\,\Omega(\mathbf{x})\,d^3\mathbf{x}=\tfrac{1}{4}\beta^2. \quad . \quad . \quad . \quad . \quad (5.20)$$

Hence with (5.17)

$$J=-\frac{5}{16}\alpha\beta+\frac{1}{4}\beta^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.21)$$

It will be noticed that the wave functions used here are all normalized so that from (5.16),

$$\beta=\frac{5}{8}\alpha, \qquad \text{and hence} \qquad J=-\frac{25}{256}\alpha^2. \quad . \quad . \quad . \quad (5.22)$$

In ordinary units, therefore, with the help of (2.37)–(2.39),

$$E(0)=\hbar\omega J=-\frac{25}{256}\alpha^2\hbar\omega=-\frac{25}{512}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right)^2\frac{e^4m}{\hbar^2}. \quad . \quad (5.23)$$

The polarization potential $\Phi_{ir}(\mathbf{y})$ at a position $\mathbf{y}$ defined by (2.30) and (4.54) is an operator whose essential part is now, however, a c-number. For from (5.10) and (5.11),

$$b_\nu^+=i(4\pi\alpha/S)^{1/2}(1/v)\int_S \exp(i\nu\mathbf{x})\Omega^+(\mathbf{x})\Omega(\mathbf{x})\,d^3\mathbf{k}+d_\nu^+, \quad . \quad (5.24)$$

where d_ν and d_ν^+ are operators satisfying a harmonic oscillator equation

$$d_\nu^+ d_\nu\,|\,n_\nu\rangle=n_\nu\,|\,n_\nu\rangle \quad \text{i.e.} \quad \langle n_\nu\,|\,d_\nu^+\,|\,n_\nu\rangle=0. \quad . \quad . \quad (5.25)$$

Hence from (4.54) and (5.24)

$$e\Phi_{ir}(\mathbf{y})=-(4\pi\alpha/S)\hbar\omega\sum_\nu(1/v^2)\int_S d^3\mathbf{x}\exp(i\nu\,.\,\mathbf{x}-\mathbf{y})\Omega^+(\mathbf{x})\Omega(\mathbf{x})$$
$$+\text{c.c.}+e\Phi_{ir}' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.26)$$

where Φ_{ir}' depends on the d_ν's in such a way that its expectation value vanishes. The main part is a c-number and is equal to expression (5.13) (cf. (5.12), replacing there $\mathbf{x}$ by $\mathbf{y}$ and $\mathbf{x}'$ by $\mathbf{x}$). Hence $\Phi_{ir}(\mathbf{y})$ is the polarization potential of the static charge distribution $\Omega^+(\mathbf{x})\Omega(\mathbf{x})$, and the polarization $\mathbf{P}_{ir}$ can similarly be derived from the electric displacement $\mathbf{D}$ due to the charge distribution $\Omega^+(\mathbf{x})\Omega(\mathbf{x})$. Thus from (5.26) is is found that (c.c. leads to a factor two)

$$\sum_1^3 \frac{\partial^2}{\partial y_i^2}\Phi_{ir}(\mathbf{y})=\frac{2}{e}4\pi\alpha\hbar\omega\,\Omega^+(\mathbf{x})\Omega(\mathbf{x}), \quad . \quad . \quad . \quad (5.27)$$

or in ordinary units using (2.37)–(2.39),

$$\nabla_{\mathbf{r}}{}^2\Phi_{ir}(\mathbf{r})=4\pi\rho(|\,\mathbf{r}-\mathbf{r}_0\,|)\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right), \quad . \quad . \quad . \quad . \quad (5.28)$$

where

$$\rho(|\,\mathbf{r}-\mathbf{r}_0\,|)=eu^3\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right)\Omega^+(|\,\mathbf{x}-\mathbf{x}_0\,|)\Omega(|\,\mathbf{x}-\mathbf{x}_0\,|) \quad . \quad . \quad (5.29)$$

is the electric charge density. Spherical symmetry of both Φ_{ir} and ρ

around the point $\mathbf{r}_0=\mathbf{x}_0/u$ has been indicated. Clearly using (2.30) and

$$\operatorname{div}\mathbf{D}=4\pi\rho(|\mathbf{r}-\mathbf{r}_0|). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.30)$$

one finds with the appropriate boundary condition

$$\mathbf{P}_{ir}=\mathbf{D}/\gamma\omega^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.31)$$

using the definition (2.14).

This shows that the treatment of the polarization field in the present section is essentially a classical one based on the static charge distribution ρ. In fact in this case (5.31) follows immediately from (2.8) when $\dot{\mathbf{P}}_{ir}=0$.

This classical approach can be used to estimate the effective mass of the system in the present model (Landau and Pekar 1948, Pelzer 1950). For this purpose the electron will be considered as a rigid charge distribution moving with a small velocity. This means that $\mathbf{r}_0$ is now a point moving through the field, i.e.

$$\mathbf{r}_0=\dot{\mathbf{r}}_0 t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.32)$$

where $\dot{\mathbf{r}}_0$ is the constant velocity. Since now $\dot{\mathbf{P}}_{ir}$ no longer vanishes, (5.31) has to be replaced by (2.8), or (5.28) by

$$\nabla_{\mathbf{r}}{}^2(\ddot{\Phi}_{ir}/\omega^2+\Phi_{ir})=4\pi\rho(|\mathbf{r}-\dot{\mathbf{r}}_0 t|)\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right). \quad . \quad . \quad . \quad (5.33)$$

Now if $\dot{\mathbf{r}}_0$ is sufficiently small then

$$\Phi_{ir}=\Phi_0(|\mathbf{r}-\dot{\mathbf{r}}_0 t|)+\Phi_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.34)$$

where Φ_0 satisfies (5.28) with $\mathbf{r}_0$ from (5.32), and $|\Phi_1|\ll|\Phi_0|$. With this latter condition, and with

$$\dot{\Phi}_0(|\mathbf{r}-\dot{\mathbf{r}}_0 t|)=\Phi_0'\dot{\mathbf{r}}_0\,; \qquad \Phi_0'=\partial\Phi_0/\partial\mathbf{r}, \quad . \quad . \quad . \quad (5.35)$$

eqn. (5.33) leads immediately to

$$\Phi_1=-(1/\omega^2)\ddot{\Phi}_0=-(\mathbf{r}_0{}^2/\omega^2)\Phi_0''. \quad . \quad . \quad . \quad . \quad (5.36)$$

In this perturbation treatment the total energy of the system is equal to the energy in the static case, (5.23), augmented by terms proportional to $\dot{\mathbf{r}}_0{}^2$. These terms can be obtained from the classical Hamiltonian (2.22). Clearly from (2.30) and (5.34),

$$4\pi\mathbf{P}_{ir}=\operatorname{grad}\Phi_0-(\dot{\mathbf{r}}_0{}^2/\omega^2)\operatorname{grad}\Phi_0'', \quad . \quad . \quad . \quad (5.37)$$

and using (5.35)

$$4\pi\dot{\mathbf{P}}_{ir}=\dot{\mathbf{r}}_0\operatorname{grad}\Phi_0'-(\mathbf{r}_0{}^3/\omega^2)\operatorname{grad}\Phi_0'''. \quad . \quad . \quad . \quad (5.38)$$

Thus since with (5.30) and (2.30)

$$-\int\mathbf{D}\mathbf{P}_{ir}\,d^3\mathbf{r}=\int\rho\Phi_{ir}\,d^3\mathbf{r}, \quad . \quad . \quad . \quad . \quad . \quad (5.39)$$

the terms of (2.22) which are proportional to $\dot{\mathbf{r}}_0{}^2$ are given by (definition of m^*)

$$\tfrac{1}{2}m^*\dot{\mathbf{r}}_0{}^2=\tfrac{1}{2}m\dot{\mathbf{r}}_0{}^2+\tfrac{1}{2}(\gamma/4\pi^2)\dot{\mathbf{r}}_0{}^2\int\{(\operatorname{grad}\Phi_0')^2-2\operatorname{grad}\Phi_0\operatorname{grad}\Phi_0''\}$$
$$-(\dot{\mathbf{r}}_0{}^2/\omega^2)\int\rho\Phi_0''. \quad . \quad . \quad (5.40)$$

A term $\frac{1}{2}m\dot{\mathbf{r}}_0{}^2$ for the additional kinetic energy of the electron has been added.

Both ρ and Φ_0 depend on $|\mathbf{r}_0-\dot{\mathbf{r}}_0 t|$ only, and the integration is straightforward. Integration by parts and use of (2.14) and (5.28) yields

$$\tfrac{1}{2}m^*=\tfrac{1}{2}m+\tfrac{1}{2}(4\pi/\omega^2)(1/\epsilon_\infty-1/\epsilon)\int\rho^2(\mathbf{r})\,d^3\mathbf{r}. \quad . \quad . \quad (5.41)$$

From (5.29), (5.14)

$$\int\rho^2(\mathbf{r})\,d^3\mathbf{r}=4\pi e^2u^3(\beta^3/8\pi)^2\int_0^\infty\exp(-2\beta x)\,x^2\,dx=e^2u^3\beta^3/64\pi, \quad (5.42)$$

so that

$$\frac{m^*}{m}=1+\frac{e^2u^3\beta^3}{8m\omega^2}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.43)$$

Introducing β from (5.22), and using (2.37)–(2.39), finally

$$m^*/m=1+\tfrac{1}{2}(5/8)^3\alpha^4. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.44)$$

§ 6. Discussion of Solutions and of Methods

6.1. *Results*

The solutions for the lowest state of an electron in a polar crystal having a self energy $E(0)$ were derived in §§ 4 and 5 with the help of the variational method. Since the energy values obtained in this method are above the exact value (or equal to it) it follows that the lower of the two solutions (4.50) and (5.23) is the better one. Hence

$$E(0)=-\alpha\hbar\omega=-\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right)e^2\left(\frac{m\omega}{2\hbar}\right)^{1/2} \quad \text{if} \quad \alpha<\frac{256}{25}, \quad (6.1)$$

and

$$E(0)=-\frac{25}{256}\alpha^2\hbar\omega=-\frac{25}{256}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon}\right)\frac{e^4m}{\hbar^2} \quad \text{if} \quad \alpha>\frac{256}{25}. \quad (6.2)$$

In order to see which of the two solutions holds in practical cases consider sodium chloride (NaCl). Here approximately $\epsilon=5{\cdot}6$ and $\epsilon_\infty=2{\cdot}3$. The wavelength of the residual rays is 61 μ and the corresponding frequency is $\omega_r/2\pi\simeq\frac{1}{2}\times10^{13}$ sec^{-1}. Hence with eqn. (2.10), $\omega\simeq4{\cdot}9\times10^{13}$ sec^{-1}. Using for m the electronic mass, $m=m_{el}$, eqns. (2.37) and (2.38) lead to $\alpha\simeq6$, $u\simeq10^7$ cm^{-1}, $\hbar\omega\simeq0{\cdot}03$ ev. It follows then that (6.1) applies to this case leading to a self energy $E(0)\simeq0{\cdot}18$ ev and to an effective mass (cf. (4.51)) of $m^*=m\ (1+\alpha/6)\simeq2m$. It will be noticed from (2.37) and (2.38) that $\alpha\propto1/\sqrt{\omega}$. Assuming $m=m_{el}$, this is the only factor which may vary substantially from substance to substance. In NaCl, the frequency ω is rather small so that for most substances smaller α-values than $\alpha=6$ should be expected, i.e. in general (6.1) should hold and m^*/m should be nearly of order unity. The properties of an electron in such a substance are then not changed substantially through its interaction with the polarization field, at any rate as long as its energy is less than $\hbar\omega$ above the ground state (cf. (4.52)). It is also seen that $u^{-1}\simeq10^{-7}$ cm is a length sufficiently larger than the lattice distance which justifies the approximation of the lattice as a continuum.

It is not easy to find actual substances for which α is so large as to make (6.2) the better solution. The possibility of large α might occur in conduction in impurity bands where m/m_{el} may be rather large, since $\alpha \propto (m/m_{el})^{1/2}$. Values of the effective mass of the order $100m_{el}$ are indicated by the work of Breckenridge and Hosler (1953) on titanium dioxide semiconductors, but even there it will be seen that the method of § 4 gives better results. In this substance roughly $\epsilon \sim 110$, $\epsilon_\infty \sim 8$, $\omega \sim 10^{14}\,\mathrm{sec}^{-1}$ and hence $u \sim 1{\cdot}7 \times 10^7\,(m/m_{el})^{1/2}\,\mathrm{cm}^{-1}$ and $\alpha \sim 2(m/m_{el})^{1/2}$; the estimate is very crude in particular since TiO_2 is not isotropic. To make here $\alpha = 256/25 \simeq 10$ requires $m/m_{el} \sim 25$. Hence from (4.51), $m^*/m = 1+\alpha/6 \simeq 3$, or $m^*/m_{el} \sim 75$ which is of the correct order of magnitude. The method of § 4 might, therefore, in this case just have reached its limit. To apply the method of § 5 would require $\alpha \geqslant 10$, i.e. $m/m_{el} \gtrsim 25$. In this case according to (5.44), $m^*/m = 1+\frac{1}{2}(5/8)^3\alpha^4 \gtrsim 1200$, and hence $m^*/m_{el} \gtrsim 30\,000$, a value which is much too large.

6.2. *Method*

Clearly it seems desirable to develop a method which leads to a continuous transition between the two methods, i.e. in which the energy $E(0)$ is of the form $-\hbar\omega(c_1\alpha+c_2\alpha^2)$. For this purpose a variational method which employs a combination of the wave functions used in §§ 4 and 5 might be successful. Buckingham (1954) has in fact carried through such a calculation. The transition has been found, however, not to be continuous ; up to a value α_0 which is slightly smaller than 256/25 it was found that the results of § 4 are exactly reproduced. At $\alpha = \alpha_0$ there is an abrupt change in wave function and a discontinuity of $\partial E(0)/\partial\alpha$. At $\alpha > \alpha_0$ the results obtained in § 5 are slightly modified. These unsatisfactory features are I think largely due to some unsatisfactory properties of the method used in § 5 connected with the concentration of the wave functions (5.1), (5.14) around an arbitrary point $\mathbf{x}_0$ in space. The wave functions (4.29), (4.26) of § 4 do not have this disadvantage. For on replacing there $\mathbf{x}$ in the wave function by $\mathbf{x}-\mathbf{x}_1$, where $\mathbf{x}_1$ is a constant vector, the wave function (4.29) is simply multiplied by a phase factor $\exp(i\mathbf{k}\mathbf{x}_1)$. To show this it must be proved that the functions $\psi(\mathbf{k}, \mathbf{v})$ given by (4.26) are not changed when replacing there $\mathbf{x}$ by $\mathbf{x}-\mathbf{x}_1$ (without altering $\mathbf{x}$ in the Hamiltonian F). This means that in (4.26) $c(k, v, \xi)\exp(-i\mathbf{v}\mathbf{x})$ must be replaced by $c_1(k, v, \xi)\exp(-i\mathbf{v}\,.\,\mathbf{x}-\mathbf{x}_1)$ where the new coefficients $c_1(k, v, \xi)$ must be determined by the variational procedure. It is then found immediately that

$$c(k, v, \xi) = c_1(k, v, \xi)\exp(i\mathbf{v}\mathbf{x}_1), \quad . \quad . \quad . \quad . \quad . \quad (6.3)$$

so that $\psi(\mathbf{k}, \mathbf{v})$ remains unaltered, as required.

The property of the total wave function Ψ, discussed above, namely

$$\Psi(\mathbf{x}-\mathbf{x}_1) = \exp(-i\mathbf{k}\mathbf{x}_1)\Psi(\mathbf{x}), \quad . \quad . \quad . \quad . \quad . \quad (6.4)$$

where $\mathbf{k}$ is the total wave vector can also be derived directly from the

Hamiltonian F as has been shown by Zienau (1954). It is closely connected with the conservation of the total wave vector. The wave functions used in § 5 do not satisfy this requirement. Nevertheless they can of course be used as trial wave functions in a variational procedure, and in fact as has been shown, for very strong interaction ($\alpha > 256/25$) they lead to better energies than the wave functions used in § 4. One might think of modifying the trial wave functions used in § 5 so as to satisfy the requirement (6.4). This is not possible, however, without completely changing the method of the self consistent field used in § 5. The difficulty is closely connected with the fact that in this method the total wave vector is not on principal axes. Thus with its definition by (3.8), using the wave function (5.1),

$$\left(-i\frac{\partial}{\partial \mathbf{x}}+\sum_{\nu}\mathbf{\nu} b_{\nu}{}^{+}b_{\nu}\right)\Omega(\mathbf{x})\Phi \neq \mathbf{k}\Omega(\mathbf{x})\Phi, \quad . \quad . \quad . \quad . \quad (6.5)$$

where $\mathbf{k}$ is a constant vector ; with the wave function (4.29), however,

$$\left(-i\frac{\partial}{\partial \mathbf{x}}+\sum_{\nu}\mathbf{\nu} b_{\nu}{}^{+}b_{\nu}\right)\exp(i\mathbf{kx})\prod_{\nu}\psi(\mathbf{k},\mathbf{\nu})\chi=\mathbf{k}\exp(i\mathbf{kx})\prod_{\nu}\psi(\mathbf{k},\mathbf{\nu})\chi, \quad . \quad (6.6)$$

as has been demonstrated with the help of (4.30).

The difficulty (6.5) encountered in the self consistent method in § 5 arises always in the use of the Hartree method, e.g. when solving the problem of n interacting particles in terms of a product of n wave functions each depending on the coordinates of a single particle (or an appropriately symmetrized combination of such wave functions). This can be demonstrated very simply in the case of two particles, e.g. a proton (coordinates $\mathbf{x}_p$) and an electron (coordinates $\mathbf{x}_e$) forming a hydrogen atom. The exact solution is of the form

$$\psi=\exp(i\mathbf{kx}_c)\phi(\mathbf{x}_e-\mathbf{x}_p), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.7)$$

where $\mathbf{x}_c$ represents the centre of gravity coordinates, and $(\mathbf{x}_e-\mathbf{x}_p)$ the relative coordinates. The total momentum operator is given by $-i\hbar\partial/\partial\mathbf{x}_c$ ($=\hbar\times$ total wave vector) and is on principal axes. In the ground state $\phi(\mathbf{x}_e-\mathbf{x}_p)$ is a function which is large only near $\mathbf{x}_e-\mathbf{x}_p=0$. It means that the electron is found always near the proton, but this may be anywhere in the total space available.

In the Hartree approximation, on the other hand, one uses a trial function

$$\psi_H=\psi_e(\mathbf{x}_e)\psi_p(\mathbf{x}_p), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.8)$$

where the functions $\psi_e(\mathbf{x}_e)$ and $\psi_p(\mathbf{x}_p)$ are to be determined from the variational method. This leads to a Schrödinger equation each for the electron and for the proton. The respective potential energies are obtained from the potentials of the average charge distributions of the proton and the electron, respectively. In order to obtain a satisfactory result for the ground state, both ψ_e and ψ_p must be spherically symmetrical functions centred at an arbitrary point $\mathbf{x}_0$ in space and being large in the neighbour-

hood of $\mathbf{x}_0$ only (ψ_e in a radius of the order of the Bohr radius, ψ_p in a correspondingly smaller radius). The energy of the ground state obtained in this way is a very good approximation to the correct value. The wave function of the whole system is, however, concentrated around an arbitrary point $\mathbf{x}_0$. In the exact solution the function $\phi(\mathbf{x}_e-\mathbf{x}_p)$ (6.7) is of course also concentrated around the point $\mathbf{x}_e-\mathbf{x}_p=0$; the function $\exp(i\mathbf{k}\mathbf{x}_c)$ provides, however, for the atom to be with equal probability at any point in space. In the Hartree wave function (6.8) it would not be permissible to multiply ψ_H by $\exp(i\mathbf{k}\mathbf{x}_0)$ and to use $\mathbf{x}_0$ as a further available variable coordinate vector because all six degrees of freedom have already been used in $\psi_e(\mathbf{x}_e)$ and $\psi_p(\mathbf{x}_p)$. It is seen clearly now how this is connected with the fact that $\partial\,\partial\mathbf{x}_c$ is now not on principal axes. This condition would restrict the relative motion of the electron and the proton just in such a way as to make three degrees of freedom available for the motion of the centre of gravity. In the Hartree approximation, however, there is no correlation between the motion of the proton and the electron, except on an average, binding both to the same point $\mathbf{x}_0$.

The wave functions used in § 5 correspond to the case discussed above. Thus both, electron wave functions $\Omega(\mathbf{x})$ and polarization potential Φ_{ir} are centred around the same point $\mathbf{x}_0$ as follows from eqns. (5.14) and (5.28).

A further difference in the use of the methods of §§ 4 and 5 can be seen in the behaviour of the polarization potential Φ_{ir}. In the Hartree method (cf. (5.28)) Φ_{ir} is essentially the classical potential of a static charge ρ. The value obtained in § 4, however, depends on the dynamic properties of the lattice (cf. (4.61), (4.62)). As a consequence the value of $E(0)$ according to the method of § 4, i.e. (6.1), depends on the dynamic properties of the field, i.e. on ω, whereas (6.2) found from the method of § 5 is independent of ω.

A method of assessing the range of validity of the method used in § 4 has been developed by Lee, Low and Pines (1953), and independently by Zienau (1953). Essentially these authors find a Hamiltonian which is satisfied exactly by the wave functions (4.29) so that one can investigate the influence of the terms which have been neglected. Gurari (1953) had already indicated this possibility. The method applied is to transform the variables (operators) O occurring in the Hamiltonian F (2.40) in such a way,

$$O'=S_t O S_t^{-1}, \quad \ldots\ldots\ldots \quad (6.9)$$

that in the new variables O' the Hamiltonian F does no longer contain terms which are linear in the field variables. In (6.9) the operator S_t must be unitary.

It is of advantage to carry through the transformation in two stages, $S_t=S_2S_1$, and choose

$$S_1=\exp\left(-i\sum_{\nu}\mathbf{v}\mathbf{x}b_{\nu}^{+}b_{\nu}\right). \quad \ldots\ldots\ldots \quad (6.10)$$

This transforms the operator of the electronic wave vector, $-i\partial/\partial\mathbf{x}$, into the total wave vector for

$$-iS_1\frac{\partial}{\partial\mathbf{x}}S_1^{-1}=-i\exp\left(-i\sum_{\mathbf{v}}\mathbf{v}\mathbf{x}b_{\mathbf{v}}^{+}b_{\mathbf{v}}\right)\frac{\partial}{\partial\mathbf{x}}\exp\left(i\sum_{\mathbf{v}}\mathbf{v}\mathbf{x}b_{\mathbf{v}}^{+}b_{\mathbf{v}}\right)$$
$$=-i\frac{\partial}{\partial\mathbf{x}}+\sum\mathbf{v}b_{\mathbf{v}}^{+}b_{\mathbf{v}}=\mathbf{k}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.11)$$

By equating it to the constant vector $\mathbf{k}$ it has been indicated that it is a constant of motion as has been shown in § 3. Also

$$b_{\mathbf{v}}'=S_1b_{\mathbf{v}}S_1^{-1}=\exp\left(-i\mathbf{v}\mathbf{x}b_{\mathbf{v}}^{+}b_{\mathbf{v}}\right)b_{\mathbf{v}}\exp\left(i\mathbf{v}\mathbf{x}b_{\mathbf{v}}^{+}b_{\mathbf{v}}\right)=b_{\mathbf{v}}\exp\left(i\mathbf{v}\mathbf{x}\right), \quad . \quad (6.12)$$

because using (2.35) (n is a positive integer),

$$b_{\mathbf{v}}b_{\mathbf{v}}^{+}b_{\mathbf{v}}=(b_{\mathbf{v}}^{+}b_{\mathbf{v}}+1)b_{\mathbf{v}}\,; \quad b_{\mathbf{v}}(b_{\mathbf{v}}^{+}b_{\mathbf{v}})^n=(b_{\mathbf{v}}^{+}b_{\mathbf{v}}+1)^nb_{\mathbf{v}}, \quad . \quad (6.13)$$

and similarly for $b_{\mathbf{v}}^{+}$. Hence $b_{\mathbf{v}}^{+}b_{\mathbf{v}}=b_{\mathbf{v}}'^{+}b_{\mathbf{v}}'$. Thus inserting from (6.11) and (6.12) into expression (2.40) for F,

$$F=(\mathbf{k}-\sum_{\mathbf{v}}\mathbf{v}b_{\mathbf{v}}'^{+}b_{\mathbf{v}}')^2+\sum_{\mathbf{v}}b_{\mathbf{v}}'^{+}b_{\mathbf{v}}'+i(4\pi\alpha/S)^{1/2}\sum_{\mathbf{v}}(b_{\mathbf{v}}'^{+}-b_{\mathbf{v}}')/v. \quad . \quad (6.14)$$

In this Hamiltonian the total wave vector of the system is $\mathbf{k}$, but $\mathbf{k}$ may have any value. In this way the number of available variables has been reduced by three in a similar way as the number of available variables of a system of interacting particles can be reduced by three by eliminating the centre of gravity motion.

Now choose

$$S_2=\exp\left[-\sum_{\mathbf{v}}c(k,v,\xi)(b_{\mathbf{v}}'^{+}+b_{\mathbf{v}}')\right], \quad . \quad . \quad . \quad . \quad (6.15)$$

and let the (imaginary) constants $c(k,v,\xi)$ be given by (4.41). Hence

$$\beta_{\mathbf{v}}=S_2b_{\mathbf{v}}'S_2^{-1}=\exp\left[c(k,v,\xi)(b_{\mathbf{v}}'^{+}+b_{\mathbf{v}}')\right]b_{\mathbf{v}}'\exp\left[c^*(k,v.\,\xi)(b_{\mathbf{v}}'+b_{\mathbf{v}}'^{-})\right]$$
$$=b_{\mathbf{v}}'+c^*(k,v,\xi), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.16)$$

since

$$c(k,v,\xi)=-c^*(k,v,\xi) \quad \text{and} \quad [b_{\mathbf{v}}',b_{\mathbf{v}}'^{+n}]=nb_{\mathbf{v}}'^{+n-1}. \quad . \quad (6.17)$$

It will be seen that with the choice (4.41) for $c(k,v,\xi)$ the linear terms in $\beta_{\mathbf{v}}$ just cancel when (6.16) is inserted into (6.14). Making, therefore, use of the definition (4.43), (4.44) of $E(\mathbf{k})$, it is found that

$$F=(E(\mathbf{k})/\hbar\omega)-k^2+(\mathbf{k}-\sum\mathbf{v}\beta_{\mathbf{v}}^{+}\beta_{\mathbf{v}})^2+F_1+F_2, \quad . \quad . \quad (6.18)$$

where

$$F_1=S^{-1}\sum_{\mathbf{v},\mathbf{v}'}(\mathbf{v}\,.\,\mathbf{v}')c(k,v,\xi)c^*(k,v',\xi')(2\beta_{\mathbf{v}}^{+}\beta_{\mathbf{v}'}-\beta_{\mathbf{v}}^{+}\beta_{\mathbf{v}'}^{+}-\beta_{\mathbf{v}}\beta_{\mathbf{v}'}). \quad (6.19)$$

and

$$F_2=-2S^{-1/2}\sum_{\mathbf{v},\mathbf{v}'}(c(k,v,\xi)\beta_{\mathbf{v}}^{+}\beta_{\mathbf{v}'}^{+}\beta_{\mathbf{v}'}+c^*(k,v,\xi)\beta_{\mathbf{v}'}^{+}\beta_{\mathbf{v}}\beta_{\mathbf{v}}). \quad (6.20)$$

The wave function now depends on the variables $\beta_{\mathbf{v}}$, $\beta_{\mathbf{v}}^{+}$. In particular, it is seen that a product of oscillator wave functions $|\,n_{\mathbf{v}}\rangle$ for the oscillators $\beta_{\mathbf{v}}^{+}\beta_{\mathbf{v}}$ satisfies

$$\{E(\mathbf{k})/\hbar\omega-k^2+(\mathbf{k}-\sum_{\mathbf{v}}\mathbf{v}\beta_{\mathbf{v}}^{+}\beta_{\mathbf{v}})^2\}\prod_{\mathbf{v}}|\,n_{\mathbf{v}}\rangle$$
$$=\{E(\mathbf{k})/\hbar\omega-k^2+(\mathbf{k}-\sum_{\mathbf{v}}\mathbf{v}n_{\mathbf{v}})^2\}\prod_{\mathbf{v}}|\,n_{\mathbf{v}}\rangle, \quad . \quad . \quad . \quad . \quad (6.21)$$

and is, therefore, a solution for the Hamiltonian $F-F_1-F_2$. Hence the wave function χ, cf. (4.8), satisfies

$$(F-F_1-F_2)\chi=(E(\mathbf{k})/\hbar\omega)\chi \quad . \quad . \quad . \quad . \quad . \quad (6.22)$$

i.e. the solution of § 4 for a given $\mathbf{k}$ is identical with the lowest solution of $F-F_1-F_2$.

To estimate the error committed in this approximation one can treat F_1+F_2 as a perturbation. This will not be carried out here in detail. By second order perturbation including some numerical work, Lee, Low and Pines (1953) find a correction

$$\Delta E=-(0{\cdot}007\alpha^2+0{\cdot}01k^2\alpha^2/(1+\alpha/6)^2)\hbar\omega. \quad . \quad . \quad . \quad (6.23)$$

Thus for $\mathbf{k}=0$, one finds α^2 terms as in § 5, though with a much smaller factor. This means that when $\alpha=256{,}25$, perturbation theory for F_1+F_2 is no longer valid. The range of validity of perturbation theory for F_1+F_2 should in fact be restricted to α-values for which

$$|\Delta E| \ll \hbar\omega, \quad \text{i.e.} \quad \alpha^2 \ll 160. \quad . \quad . \quad . \quad . \quad . \quad (6.24)$$

At $\alpha=256{,}25\simeq 10$. $\alpha^2\simeq 100$ which is not very small compared with 160. For $\alpha=6$, however the value for NaCl, $\alpha^2=36$, the correction (6.23) is sufficiently smaller than $\hbar\omega$ to make the results of § 4 reliable.

Equation (6.18) also shows that the case $E(\mathbf{k})-E(0)>\hbar\omega$ leads to difficulties because then the solutions of $F-F_1-F_2$ are then degenerate.

In conclusion, therefore, it has been found that the method of § 4 gives reliable solutions for energy and effective mass if the interaction parameter α satisfies $\alpha^2\ll 100$ and provided $E(\mathbf{k})-E(0)<\hbar\omega$. There is no obvious way in which this method can be extended beyond these limits. The method of § 5 should give reasonable energies $E(0)$ for $\alpha^2>100$ though in this case it may no longer be justifiable to treat the lattice as a continuum. This method leads to difficulties inherent in the Hartree method and connected with the motion of the centre of gravity of the system. The validity of the calculation of the effective mass in this method is, therefore, unclear.

It should finally be remarked that application of the method of § 5 to metals has led to an understanding of the superconductive properties of a one dimensional superconductor (Fröhlich 1954) but does not yield the isotope effect which is obtained, however, by methods corresponding to those used in § 4. This agrees with the observation that the method of § 5 does not introduce the dynamic properties of the lattice. Thus in this case too development of a method linking the results of § 4 and § 5 would be very desirable.

§ 7. Mobility

If a weak external electric field is applied to the system consisting of an electron and the polarization field then an electric current is established. Its magnitude can be described in terms of the microscopic mobility μ

which is the average electronic velocity per unit field strength. By definition μ is connected with the average relaxation time $\bar{\tau}$ by

$$\mu = e\bar{\tau}/m. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.1)$$

The methods for calculating μ or $\bar{\tau}$ are well established in the case of sufficiently weak interaction between electron and lattice vibrations, i.e. when perturbation theory holds ($\alpha < 1$). In this case the electron (or electrons) is considered as nearly independent of the polarization quanta. The electron is then accelerated by the field, and is scattered by occasional absorption or emission of a quantum. Assuming the quanta to be in thermal equilibrium, the electron carries an average electric current which is proportional to the strength of the field.

This procedure cannot be generalized in a straightforward manner to the case of stronger coupling ($\alpha > 1$). For this would require the possibility of separating the system into free quanta on the one hand, and the electron plus ' bound ' quanta on the other. The methods developed in previous sections, however, involved the assumption that the total energy (above the ground state) of the whole system is less than $\hbar\omega$ which excludes the presence of free quanta. The difficulties which one meets in attempts to go to higher total energies can be seen from the transformed Hamiltonian in the form (6.18)–(6.20). There if $\alpha^2 \ll 100$, the terms $F_1 + F_2$ can be considered as negligibly small provided no ' free ' quanta are present so that the wave function is given by χ and the energy by $E(\mathbf{k})$, as shown in (6.22). If, however, $n_{\mathbf{v}}$ quanta of wave vector $\mathbf{v}$ are excited, as considered in (6.21), then the previously neglected terms $F_1 + F_2$ may become extremely large so that the separation of free quanta achieved in the Hamiltonian $F - F_1 - F_2$ has no longer any physical significance.

In view of these difficulties application of perturbation theory only will be discussed in the present section. It may be hoped that some of the results obtained in this way have a validity beyond $\alpha = 1$, although no proof for this can be given at present. This hope arises from the calculation of the energy of the ground state in § 4 according to which perturbation theory happened to give good results up to $\alpha \simeq 10$. This means, of course, that electronic energies larger than $\hbar\omega$ have to be excluded because in this energy region none of our methods could be applied. Perturbation theory does yield satisfactory results, however, for electronic energies which are very large compared with $\hbar\omega$. This case was treated by the author (Fröhlich 1937) in a paper in which the idea of polarization waves was first introduced. The probability per second for absorption or emission of a quantum by the electron was found very small compared with ω so that no objections against the use of perturbation theory arise in this case. This no longer holds, however, when the results are extrapolated to electronic energies near $\hbar\omega$ in agreement with the conclusions of the present discussion. If the electronic energy is less than $\hbar\omega$, however, then quanta can no longer be emitted

but only be absorbed. If furthermore the temperature is so low that very few quanta only are present then the probability for this process is very small. It may be hoped then that perturbation theory may be applied to this case as also suggested above. The calculation of the mobility was carried out by Fröhlich and Mott (1937) and corrected for a more correct electron lattice interaction by Fröhlich, Pelzer and Zienau (1950). In this latter paper it was indicated that after absorption of a quantum the electron has an energy larger than $\hbar\omega$ above the ground state and hence can emit a quantum. The probability for this process in general is so large that perturbation theory cannot be applied, and this is the main reason for the uncertainty of the validity of perturbation theory even in the case of low temperatures and average electronic energies below $\hbar\omega$. The fact that the probability of emission is very large suggests that the whole process of scattering should be treated similarly to that of resonance scattering of light. Thus a quantum $\hbar\omega$, wave vector $\mathbf{v}$ is absorbed by the electron and re-emitted with wave vector $\mathbf{v}'$. In this process the electron is scattered elastically, changing its wave vector by $\mathbf{v}-\mathbf{v}'$. This resonance scattering has been treated by Zienau (1950).† This process can also be treated by replacing the free electron wave function by a function of the type treated in § 4, but treating the quantum as free, as carried out by Low and Pines (1953). Their result suggests that perturbation theory does give satisfactory results in the case of low temperatures and electronic energies below $\hbar\omega$. Finally it should be mentioned that use of perturbation theory for both absorption and emission together with an appropriate treatment of the kinetic equation leads for low temperatures to the same result for the mobility as the method of resonance scattering. This follows from the paper by Howarth and Sondheimer (1953).

The above survey shows that in actual crystals where usually $1<\alpha<10$, there is some hope that application of perturbation theory for the calculation of the mobility should give reasonable results for temperatures $T\ll\hbar\omega/k$ and electronic energies sufficiently smaller than $\hbar\omega$ though as mentioned it has not been possible, so far, to prove this. In view of the theoretical uncertainties, it seems desirable not to involve the complicated theory of resonance scattering but to derive the result very simply by assuming right from the beginning that the probability per second for resonance scattering, i.e. for absorption and re-emission of a quantum is equal to that for absorption alone as would in fact follow from perturbation theory where the ratio of absorption to emission probabilities is very small, of the order $\bar{n}/(1+\bar{n})\simeq\bar{n}\ll 1$. Here

$$\bar{n}=\frac{1}{\exp(\hbar\omega/kT)-1}\ll 1\,;\quad kT\ll\hbar\omega \qquad (7.2)$$

is the average excitation probability of a quantum.

† A mistake by a factor 2/3 in the final integration has been corrected by C. Herring (1952).

To derive this probability let the electron be in a state with wave vector $\mathbf{q}$, so that its energy is $\hbar\omega q^2$ and let

$$q^2 \ll 1. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.3)$$

Also let $n_{\mathbf{v}}$ describe the number of excited quanta with wave vector $\mathbf{v}$ so that the zero order wave functions are given by (4.2). Using $\hbar\omega F_{\text{int}}$ (4.7) as perturbation, transitions into electronic states $\mathbf{q}'$ are possible in which a single quantum is absorbed, conserving energy and the total wave vector, i.e.

$$\mathbf{q}' = \mathbf{q} + \mathbf{v}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.4)$$

and

$$q'^2 = q^2 + v^2 + 2qv \cos\theta = q^2 + 1 \quad \text{or} \quad v^2 + 2vq \cos\theta = 1. \quad (7.5)$$

Here θ is the angle between $\mathbf{q}$ and $\mathbf{v}$. Since $|\cos\theta| < 1$, it follows immediately that v is restricted to the range

$$1 - q < v < 1 + q. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.6)$$

Now since

$$\langle n_{\mathbf{v}} - 1 \,|\, b_{\mathbf{v}} \,|\, n_{\mathbf{v}} \rangle = n_{\mathbf{v}}^{1/2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.7)$$

it follows from (4.7) and (4.2) that the absolute square of the matrix element for the absorption process is given by

$$|M|^2 = \frac{4\pi\alpha}{S} \frac{n_{\mathbf{v}}}{v^2} (\hbar\omega)^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.8)$$

The transition probability per second is obtained in the usual way by multiplying (7.8) by $2\pi/\hbar$ and by the number of levels in the final state per unit final energy $q'^2\hbar\omega$. Since the final electronic state is exactly given by (7.4) the density of levels is equal to the density of levels for the quanta, i.e. per unit energy using (7.5) it is given by

$$\frac{S}{(2\pi)^3} \frac{v^2\, dv\, d\cos\theta\, d\phi}{\hbar\omega\, d(q'^2)} = \frac{S}{(2\pi)^3} \frac{v^2\, dv\, d\phi}{\hbar\omega 2qv}. \quad . \quad . \quad . \quad . \quad . \quad (7.9)$$

Hence with (7.8) averaging $n_{\mathbf{v}}$ over a small $\mathbf{v}$ range,

$$dP_a = \frac{2\pi}{\hbar} \frac{4\pi\alpha}{(2\pi)^3} \frac{\hbar\omega}{2qv} \bar{n}\, dv\, d\phi \quad . \quad . \quad . \quad . \quad . \quad (7.10)$$

represents the probability of absorption of a quantum in a certain wave vector range by the electron in the state $\mathbf{q}$. The wave vector range is such that the angle θ of $\mathbf{q}$ and $\mathbf{v}$ is exactly given by (7.5), the magnitude v has a range dv, and the azimuth ϕ with $\mathbf{q}$ as axis has a range $d\phi$.

It must be noticed now that conservation of energy requires that re-emission of a quantum by the electron in the state $\mathbf{q}'$ leads to a final state $\mathbf{q}''$ such that

$$q^2 = q''^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.11)$$

The direction of the final state is arbitrary and the probability for re-emission follows from (7.10) by replacing there $\bar{n}$ by $1 + \bar{n}$. Now with (7.3) it follows from (7.6) that v has nearly the same value $v \simeq 1$ whatever

the direction of $\mathbf{q}''$ so that the probability for scattering into any of these directions is equal. Thus using (7.10), (7.6) and (7.3)

$$P_a = \int dP_a = \frac{\alpha\omega\bar{n}}{2\pi q}\int_0^{2\pi} d\phi \int_{1-q}^{1+q} \frac{dv}{v} = \frac{\alpha\omega\bar{n}}{q}\log\frac{1+q}{1-q} \simeq 2\pi\omega\bar{n}, \quad (7.12)$$

is the probability per second for any absorption process, and hence since the probability for re-emission is very much larger (provided $\bar{n} \ll 1$) P_a is also the probability per second for resonance scattering. This latter involves elastic scattering of the electron with equal probability for any scattering angle. Hence

$$P_a\, d\Omega/4\pi = 2\alpha\omega\bar{n}\, d\Omega/4\pi \quad . \quad . \quad . \quad . \quad . \quad (7.13)$$

is the probability per second for the elastic scattering of the electron in a solid angle $d\Omega$.

With the help of this transition probability the rate of change of an electronic distribution function $f(\mathbf{q})$ can be calculated. For clearly

$$\frac{\partial f(\mathbf{q})}{\partial t} = -P_a \int (f(\mathbf{q}) - f(\mathbf{q}''))\, d\Omega''/4\pi, \quad q^2 = q''^2, \quad . \quad . \quad (7.14)$$

since P_a is independent of the direction of $\mathbf{q}''$; $d\Omega''$ represents the differential of the solid angle of $\mathbf{q}''$. In thermal equilibrium, $f(\mathbf{q})$ depends on the magnitude q only and hence $\partial f/\partial t = 0$ in this case. It will be noticed, of course, that elastic scattering cannot alter the magnitude of q and hence can establish equilibrium only if $\int f(\mathbf{q})\, d\Omega$ has the equilibrium value. This is assumed to hold in the theory of conductivity in weak fields where a distribution function $f(\mathbf{q})$ of the form

$$f(\mathbf{q}) = f_0(q) + (q_F/q) f_1(q) \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.15)$$

is considered. Here q_F is the projection of $\mathbf{q}$ into the field direction, and both $f_0(q)$ and $f_1(q)$ depend on the magnitude of q only. Clearly then

$$\int (f(\mathbf{q}) - f(\mathbf{q}''))\, d\Omega'' = (f_1(q)/q) \int (q_F - q_F'')\, d\Omega''$$
$$= 4\pi(q_F/q) f_1(q) = 4\pi(f(\mathbf{q}) - f_0(q)), \quad . \quad (7.16)$$

because $\int q_F''\, d\Omega'' = 0$. From the definition of the relaxation time τ,

$$\frac{\partial f(\mathbf{q})}{\partial t} = -\frac{1}{\tau}(f(\mathbf{q}) - f_0(q)), \quad . \quad . \quad . \quad . \quad . \quad (7.17)$$

it follows with (7.16), (7.14) and (7.13) that

$$1/\tau = 2\alpha\omega\bar{n}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.18)$$

Also since τ is independent of $\mathbf{q}$, $\tau = \bar{\tau}$.

It should be remarked that second order field effects change the average electronic energy. Equilibrium can then not be established through elastic scattering processes only.

The mobility μ now follows by introducing (7.18) into (7.1). Use of (7.2) and (2.37) leads to

$$\mu = \frac{ea_0}{(2\hbar\omega m_{el})^{1/2}} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon}\right)^{-1} \left(\frac{m_{el}}{m}\right)^{3/2} [\exp(\hbar\omega/kT) - 1], \quad kT \ll \hbar\omega \quad . \quad (7.19)$$

where

$$a_0 = \hbar^2/me^2 \simeq 0{\cdot}5 \times 10^{-8} \text{ cm} \quad . \quad . \quad . \quad . \quad . \quad (7.20)$$

is the Bohr radius. Hence if

$$\Theta = \hbar\omega/k \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.21)$$

is expressed in degrees absolute,

$$\mu \simeq 10 \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon}\right)^{-1} \left(\frac{m_{el}}{m}\right)^{3/2} \frac{\exp(\Theta/T) - 1}{\sqrt{\Theta}} \text{ cm}^2 \text{ volt sec}, \quad T \ll \Theta. \quad . \quad (7.22)$$

It will be recalled that $\omega/2\pi$ is the frequency of longitudinal vibrations which is connected with the frequency of residual rays by eqn. (2.10). Hence if c is the velocity of light and λ is the residual ray wave length then from (7.21) and (2.10),

$$\Theta = \frac{\hbar}{k} \frac{2\pi c}{\lambda} \sqrt{\frac{\epsilon}{\epsilon_\infty}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.23)$$

Before comparing the theoretical value (7.22) for the microscopic mobility μ with experimental results the above mentioned difficulties in the theory must be emphasized again. Furthermore it should be remembered that there will always be contributions to the scattering of electrons due to mechanisms of a different nature from the one discussed here. Thus scattering due to the so-called acoustic vibrations will make a contribution to μ proportional to $T^{3/2}$ and should thus be decisive at very low temperatures. The magnitude of this type of scattering is in general very small, however (cf. Mayer and Polder 1953).

Experiments on NaCl have been carried out by Redfield (1953) who finds $\mu = 250 \pm 50$ cm^2/volt sec at $T = 82 \pm 2°$K and $\mu \sim 40$ at $T = 200°$K. Theoretically with $\epsilon = 5{\cdot}6$, $\epsilon_\infty = 2{\cdot}3$, $\lambda = 61 \times 10^{-4}$ cm, one has $\Theta = 340°$ and hence $\mu = 280(m_{el}/m)^{3/2}$ and $\mu = 15(m_{el}/m)^{3/2}$ respectively. The agreement is very satisfactory (with $m \sim m_{el}$). Measurements on AgCl by Haynes and Shockley (1953) can be represented by $\mu = 30[\exp(290/T) - 1]$. Theoretically with $\epsilon = 12{\cdot}3$, $\epsilon_\infty = 4{\cdot}0$ and $\Theta = 290°$, the factor 30 is replaced by $4(m_{el}/m)^{3/2}$ which has also the correct magnitude. The measured temperature range is not wide enough, however, to exclude a $T^{-3/2}$ law. Agreement with the theoretical temperature dependence over a wide range was obtained by Breckenridge and Hosler (1953) on TiO_2. Their absolute values indicate very high electronic effective mass probably due to conduction in impurity bands.

Perturbation theory which was used in the derivation of expression (7.18) for τ is also valid for electronic energies which are very high compared with $\hbar\omega$. The actual expression in this case differs, however, from (7.18) for various reasons. This case will not be considered here

because interaction with very short polarization waves is predominant in this case so that treatment of the lattice as a continuum is no longer permissible. This leads to some uncertainties concerned with the value of ω and with the correct form of the interaction, but on the other hand use of perturbation theory is well justified. In contrast to this case, however, use of perturbation theory for electronic energies of the order $\hbar\omega$ is quite unjustifiable in particular for alkali halides where $\alpha \simeq 6$.

Further experiments might help to overcome the theoretical difficulties encountered in this respect. This would require the measurement of electronic mobilities for electronic energies of the order of $\hbar\omega$ i.e. in the case of thermal equilibrium for temperatures of the order $T \simeq \Theta$.

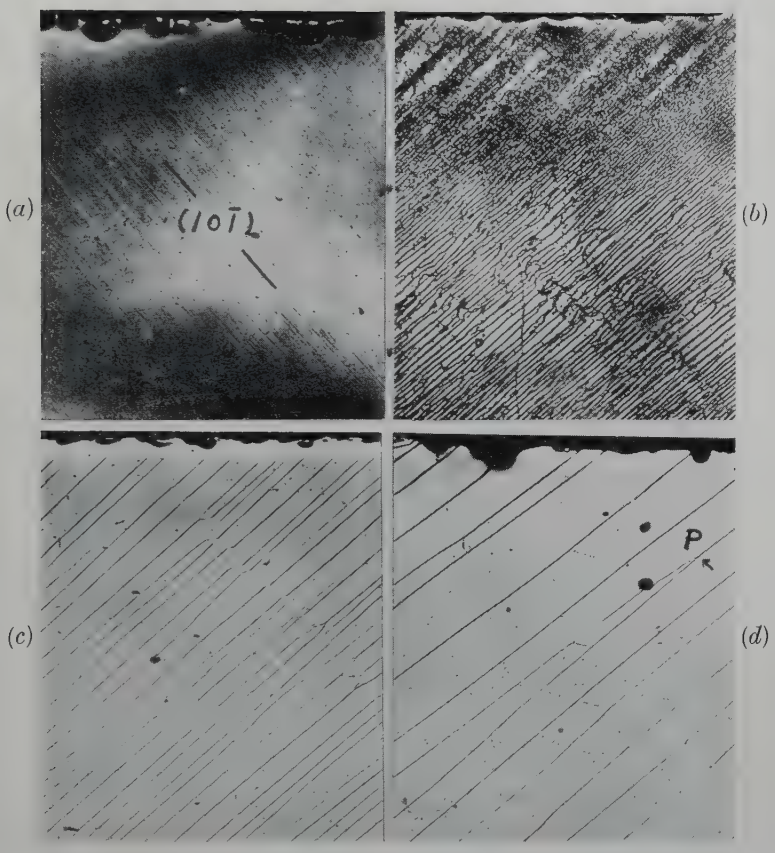
Acknowledgment

It is gratefully acknowledged that much of the work discussed here was originally carried out with the support of the British Electrical and Allied Industries Research Association (E.R.A.).

References

Bloch, F., 1928, *Z. Phys.*, **52,** 555.
Breckenridge, R., and Hosler, W., 1953, *Phys. Rev.*, **91,** 793.
Buckingham, M. J., 1954, *E.R.A. Report*, to be published.
Fröhlich, H., 1937, *Proc. Roy. Soc.* A, **160,** 280; 1949, *Theory of Dielectrics* (Oxford: Clarendon Press); 1950, *Phys. Rev.*, **79,** 845 and *Proc. Phys. Soc.* A, **63,** 778; 1953, *Physica*, **19,** 755; 1954, *Proc. Roy. Soc.* A, **223,** 296.
Fröhlich, H., and Mott, N. F., 1939, *Proc. Roy. Soc.* A, **171,** 496.
Fröhlich, H., Pelzer, H., and Zienau, S., 1950, *Phil. Mag.*, **41,** 221.
Gurari, M., 1953, *Phil. Mag.*, **44,** 329.
Haynes, J., and Shockley, W., 1951, *Phys. Rev.*, **82,** 935.
Herring, C., 1952, personal communication.
Howarth, D. J., and Sondheimer, E. H., 1953, *Proc. Roy. Soc.* A, **219,** 53.
Landau, L., 1933, *Phys. Z. Sowjetunion*, **3,** 664.
Landau, L., and Pekar, S., 1948, *Journal of Physics, U.S.S.R.*, **18,** 419.
Lee, T., and Pines, D., 1952, *Phys. Rev.*, **88,** 960.
Lee, T., Low, F., and Pines, D., 1953, *Phys. Rev.*, **90,** 297.
Low, F., and Pines, D., 1953, *Phys. Rev.*, **91,** 193.
Meyer, H., and Polder, D., 1953, *Physica*, **19,** 255.
Pekar, S., 1946, *Journal of Physics, U.S.S.R.*, **10,** 341; 1949, *Ibid.*, **19,** 796.
Pelzer, H., 1950, *E.R.A. Report* L/T243.
Redfield, A., 1953, *Phys. Rev.*, **91,** 753.
Tomonaga, S., 1947, *Progr. Theor. Phys.*, **2,** 6.
Zienau, S., 1950, *E.R.A. Report* L/T236; 1953, *E.R.A. Report* L/T299; 1954, to be published.

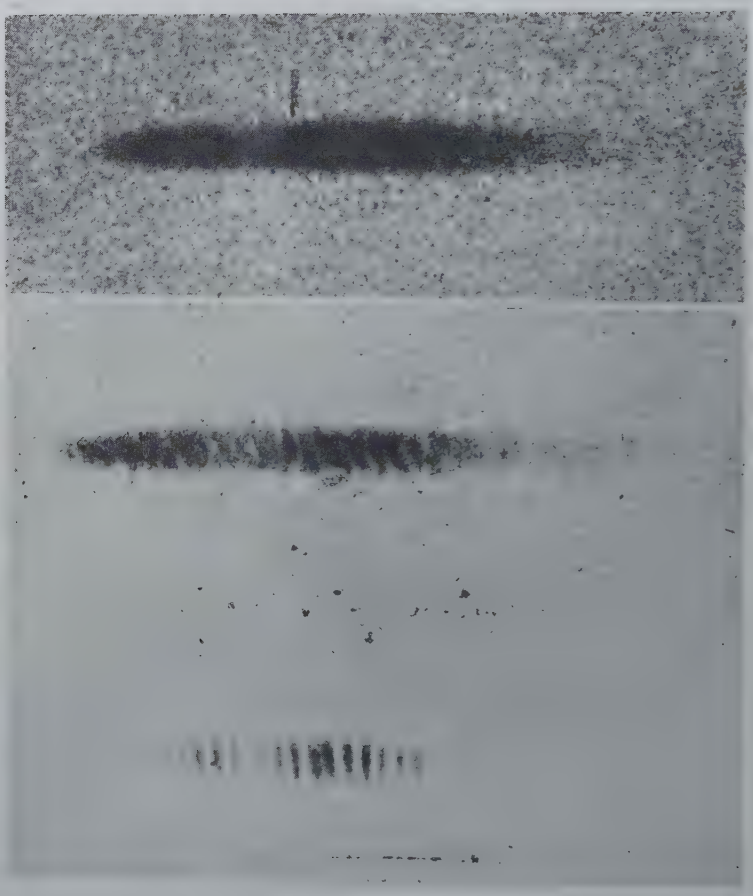
Fig. 9



Subgrain growth in bent silicon–iron single crystal as a result of various annealing treatments. [Ref. (47).]

(*a*) Bent and sectioned to show slip lines.
(*b*) Annealed 10 minutes at 950°C.
(*c*) Annealed 24 hours at 950°C.
(*d*) Annealed 4 hours at 1300°C.

Fig. 10

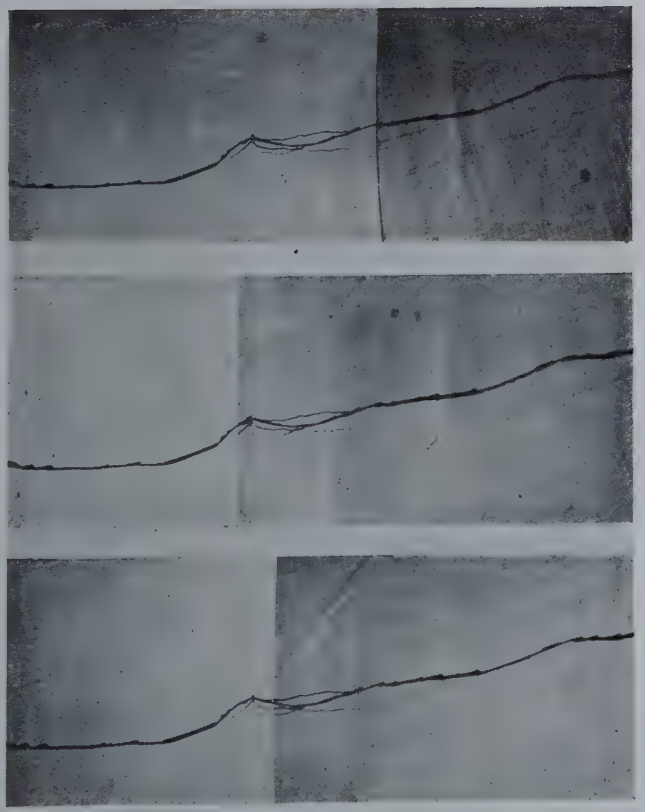


Laue asterism streak, as affected by subgrain growth, for specimens shown in fig. 9. [Ref. (47).]

Fig. 11

Subgrains in pure Al after creep, as shown with oxide film-polarized light technique. [Ref. (11).]

Fig. 12

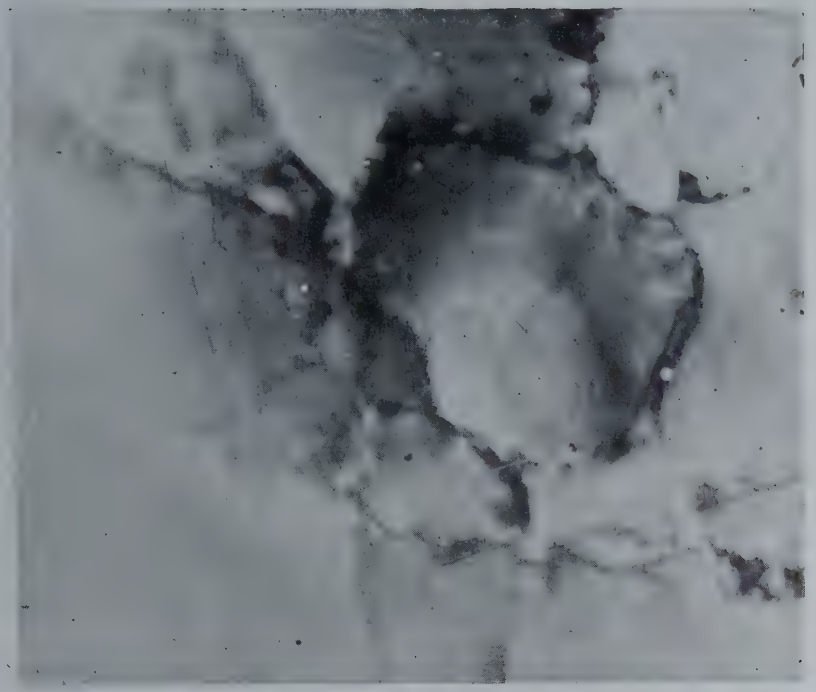


Kink boundary (2° bend) moved back and forth by the application of appropriate external shear stress, as seen on (0001) cleavage surface of pure Zn single crystal. ×50. [Ref. (52).]

Fig. 15

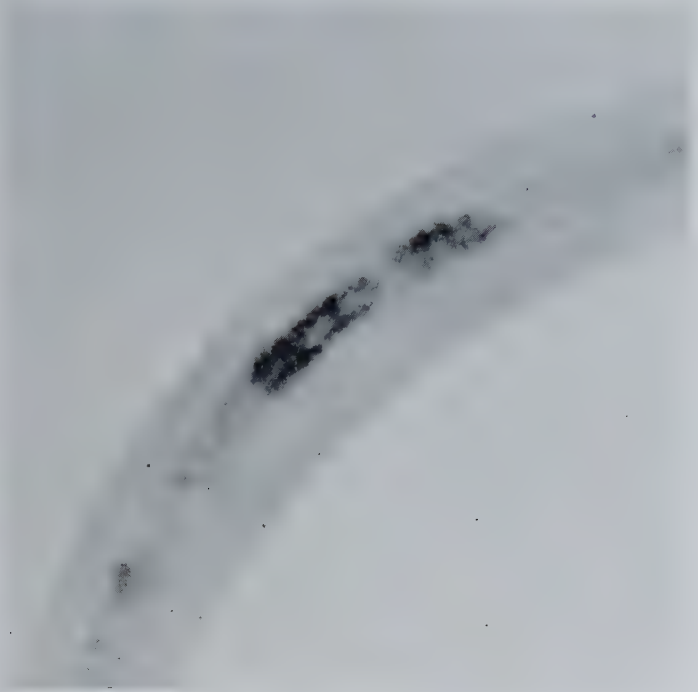
X-ray diffraction spot (222) for extended (2%) Al single crystal, unannealed, convergent monochromatic Cu Kα radiation. Streaks show subgrain formation on plastic deformation. (H. Lambot 1953.)

Fig. 16



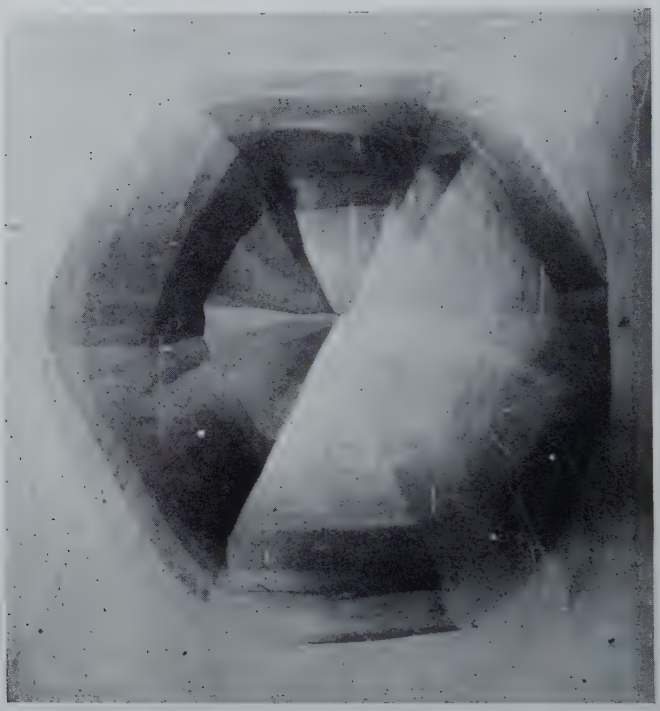
Subgrains in Al hammered at room temperature, as shown by means of transmission electron microscopy according to Heidenreich. [Ref. (71).]

Fig. 17



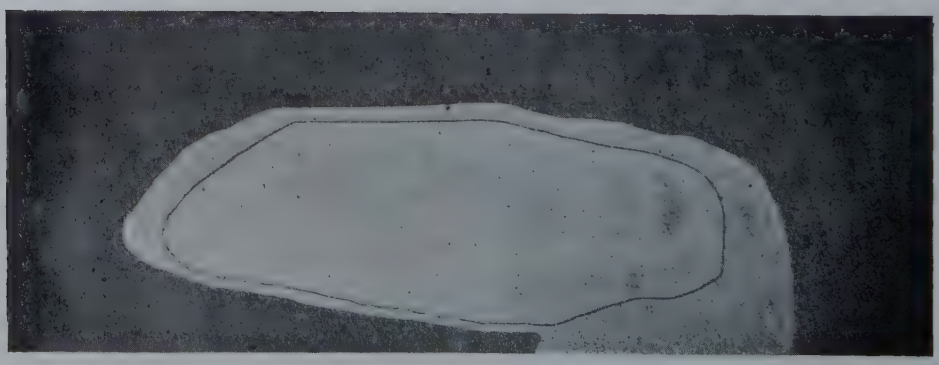
Spotty (220) x-ray diffraction ring from Fe, 50% rolled at room temperature and not annealed, using 68μ dia. microbeam, Fe, Kα radiation. (A. Kelly, Ph.D. Thesis, U. of Cambridge, 1953.)

Fig. 19



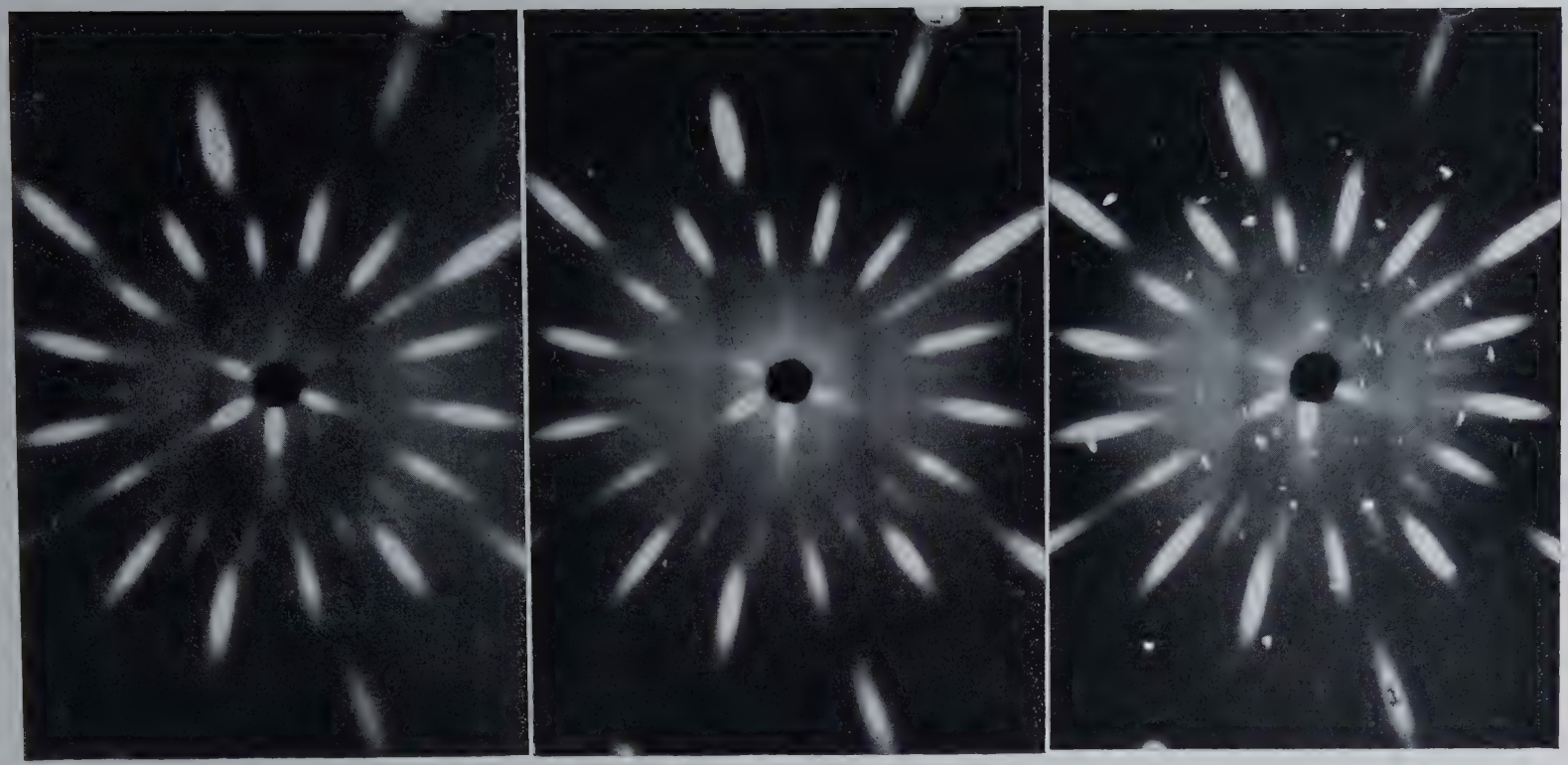
Macrophotograph of cleavage face of indented Zn crystal, as in fig. 18, showing substructure formed on indentation. [Ref. (16).]

Fig. 26



Migration of high angle grain boundary during recrystallization in pure Al. Two successive boundary positions in specimen rolled 7·5% and annealed at 500°c for 25 sec, resp. 32 sec. Oxide film and polarized light. ×75. [Ref. (11).]

Fig. 25



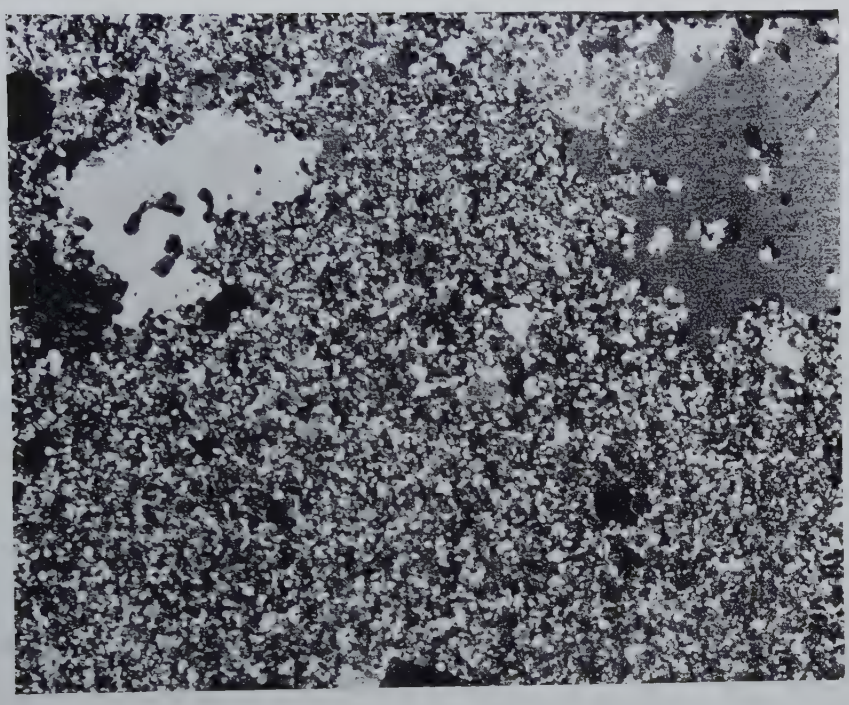
(*a*) (*b*) (*c*)

X-ray pinhole patterns of Al specimens (*a*) as rolled, (*b*) annealed 1300 seconds at 350°c and (*c*) annealed 1600 seconds at 350°c (see fig. 24). Complete softening at 1300 seconds is not accompanied by recrystallization. Partial recrystallization between 1300 and 1600 sec causes no further softening. [Ref. (30).]

Fig. 39

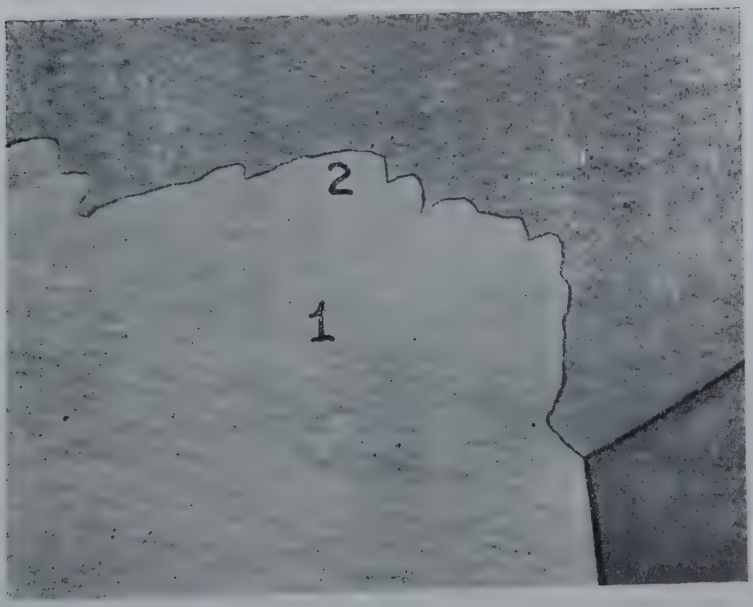
Grain boundary clinging to dispersed particles of a second phase in Al+0·74% Mn alloy after grain growth at 580°c. [Ref. (118).]

Fig. 51



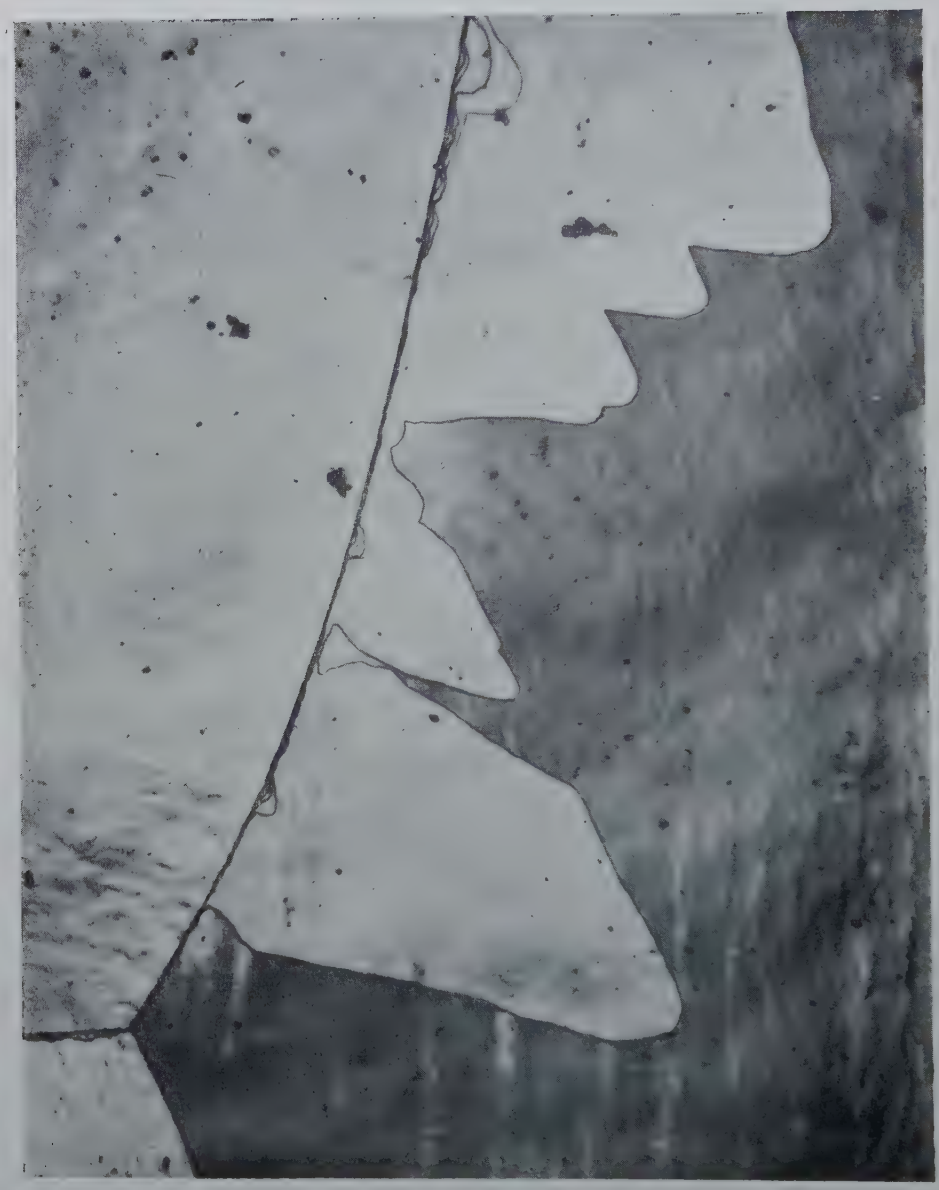
Photomicrograph showing partial coarsening in Al+1·1% Mn alloy after final anneal of 1665 minutes at 625°c. [Ref. (118).]

Fig. 57



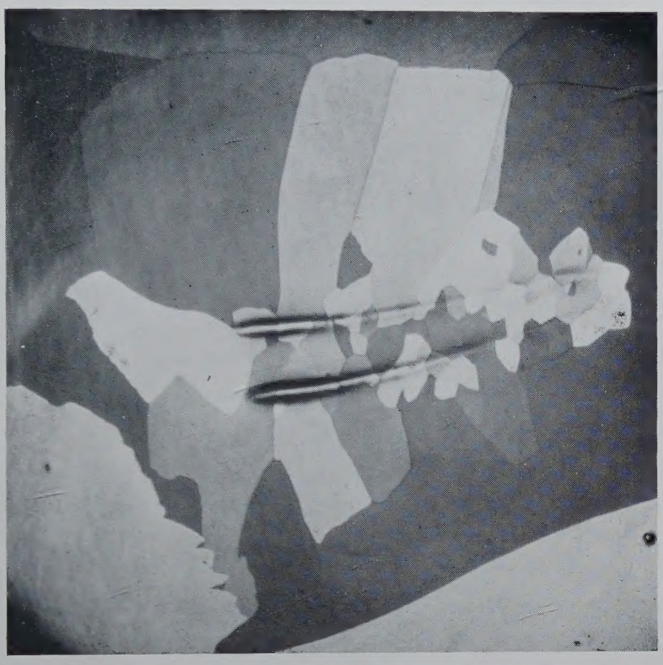
Photomicrograph showing strain-induced boundary migration in pure Al on annealing for 5 minutes at 400°c after 14% deformation by rolling. ×75. [Ref. (146).]

Fig. 58



Photomicrograph showing five successive stages of strain-induced boundary migration in pure Al on annealing at 350°c for 5, 10, 20, 30 and 150 minutes, after 12% deformation by rolling. ×75. [Ref. (11).]

Fig. 61



Photomicrograph of pure Al specimen showing many small recrystallized grains of various orientations formed in the immediate vicinity of the two surface scratches. Those recrystallized grains with favourable orientation with respect to the 12% rolled matrix grain grew selectively on annealing for 1 hour at 350°C. Electrolytic polishing, anodic etching, polarized light. ×25. [Ref. (55).]

Fig. 62

Photomicrograph showing several small recrystallized grains of varying orientations nucleated at the sheared edge. Those of the recrystallized grains with favourable orientation grew large at the expense of the annealed fine grained matrix with single orientation texture. Matrix was produced by 3 minutes anneal at 525°C of pure Al single crystal previously rolled 90%. Deep etching. ×8. [Discussion of Ref. (157).]

Annals of Science

A QUARTERLY REVIEW OF THE HISTORY OF SCIENCE SINCE THE RENAISSANCE